DOCUMENT RESUME

ED 444 367	FL 026 359

AUTHOR	Moisset, Christine, Ed.; Lipson, Mimi, Ed.
TITLE	Selected Papers from NWAVE(E) 27 (Athens, Georgia, October 1-4, 1998). University of Pennsylvania Working Papers in Linguistics, Volume 6, Number 2.
INSTITUTION	Pennsylvania Univ., Philadelphia. Penn Linguistics Club.
ISSN	ISSN-1524-9549
PUB DATE	1999-00-00
NOTE	164p.; Papers in this volume were presented at the Annual New Ways of Analyzing Variation Conference.
AVAILABLE FROM	PWPL, 619 Williams Hall, University of Pennsylvania, Philadelphia, PA 19104-6305 ($15).
PUB TYPE	Collected Works - Proceedings (021)
EDRS PRICE	MF01/PC07 Plus Postage.
DESCRIPTORS	Child Language; Creoles; Dialect Studies; Ethnic Groups; Foreign Countries; French; French Canadians; Korean; Language Minorities; *Language Research; *Linguistics; Second Language Instruction; Second Language Learning; Sociolinguistics; *Uncommonly Taught Languages
IDENTIFIERS	Canada; Korea; South Africa

ABSTRACT

This issue includes the following articles: "Vowel Epenthesis in Vimeu Picard: A Preliminary Investigation" (Julie Auger, Jeffrey Steele); "Lexical Borrowings from French in Written Quebec English: Perspectives on Motivation" (Pamela Grant-Russell and Celine Beaudet); "Variable Article Use in Korean Learners of English" (Hikyoung Lee); "The Loss of Auxiliary Selection in English" (Mimi Lipson); "Syntactic Change in Progress: Semi-Auxiliary Busy in South African English" (Rajend Mesthrie); "The Emergence of Creole Subject-Verb Agreement" (Miriam Meyerhoff); "Double Subject Marking in L2 Montreal French" (Naomi Nagy, Helene Blondeau); "Testing the Creole Continuum" (Peter Patrick); "Going Younger To Do Difference: The Role of Children in Language Change" (Julie Roberts); and "Situated Ethnicities: Constructing and Reconstructing Identity in the Sociolinguistic Interview" (Natalie Schilling-Estes). References are appended to each article. (KFT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 444 367

UNIVERSITY OF PENNSYLVANIA
WORKING PAPERS IN LINGUISTICS

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Williams, Alexander

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

FL026307

Volume 6.2 (1999): Selected Papers from NWAV(E) 27

BEST COPY AVAILABLE

UNIVERSITY OF PENNSYLVANIA
WORKING PAPERS IN LINGUISTICS

Volume 6.2 (1999)

Selected Papers from NWAV(E) 27

issue editors
Christine Moisset, Mimi Lipson

series editors
JIM ALEXANDER ALEXIS DIMITRIADIS NA-RAE HAN
ELSI KAISER MICHELLE MINNICK FOX
CHRISTINE MOISSET ALEXANDER WILLIAMS

About the PWPL series

The *University of Pennsylvania Working Papers in Linguistics* (PWPL) is an occasional series published by the Penn Linguistics Club, the graduate student organization of the Linguistics Department of the University of Pennsylvania. The series has included volumes of previously unpublished work, or work in progress, by linguists with an ongoing affiliation with the Department, as well as volumes of papers from the NWAVE conference and the Penn Linguistics Colloquium.

The current PWPL series editors are Jim Alexander, Alexis Dimitriadis, Na-Rae Han, Elsi Kaiser, Michelle Minnick Fox, Christine Moisset, and Alexander Williams.

The Editors

Publication in the *University of Pennsylvania Working Papers in Linguistics (PWPL)* does not preclude submission of papers elsewhere; copyright is retained by the author(s) of individual papers. Comments on the papers should be sent directly to the authors, at the address provided at the end of each paper.

The *PWPL* editors can be contacted at the following addresses:

U. Penn Working Papers in Linguistics
619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104–6305

working-papers@babel.ling.upenn.edu

For more information, please consult our web site:

http://ling.upenn.edu/papers/pwpl.html

Volumes of the *PWPL* are available at $15 per copy, pre-paid. (See information at the end of this volume). Multi-issue subscriptions, copies of previous issues, and standing orders (for institutions only) are also available. We invite other student organizations to set up an exchange of publications with us.

ISSN: 1524-9549

Table of Contents

About the PWPL ii

Editors' statement v

Vowel Epenthesis in Vimeu Picard: A Preliminary Investigation 1
Julie Auger and Jeffrey Steele

Lexical Borrowings from French in Written Quebec English:
Perspectives on Motivation 17
Pamela Grant-Russell and Céline Beaudet

Variable Article Use in Korean Learners of English 35
Hikyoung Lee

The Loss of Auxiliary Selection in English 49
Mimi Lipson

Syntactic Change in Progress:
Semi-auxiliary *busy* in South African English 63
Rajend Mesthrie

The Emergence of Creole Subject-Verb Agreement 77
Miriam Meyerhoff

Double Subject Marking in L2 Montreal French 93
Naomi Nagy and Hélène Blondeau

Testing the Creole Continuum 109
Peter Patrick

Going Younger to Do Difference:
The Role of Children in Language Change 121
Julie Roberts

Situated Ethnicities: Constructing and Reconstructing
Identity in the Sociolinguistic Interview 137
Natalie Schilling-Estes

Contents of Volume 6 153

From the editors

The twenty-seventh annual meeting of NWAV(E) (New Ways of Analyzing Variation) convened at the University of Georgia at Athens on October 1-4, 1998. NWAV(E) is an important forum for researchers working in the areas of sociolinguistics, dialectology, language variation, and change in general. This volume contains a selection of papers presented at NWAV(E) 27.

The papers in this volume reflect the diversity of language-related topics upon which variationist methods have been brought to bear. The authors address such social dimensions of variation as conversational style and ethnic identity (Schilling-Estes), age cohorts (Roberts), speech communities (Patrick), language contact and borrowing (Grant-Russell & Beaudet). Roberts, Lee, and Nagy & Blondeau explore the status of variable phenomena in first and second language acquisition. Explicitly linguistic subjects are also approached from a variationist perspective, including grammatical description in the areas of phonology, morphology, syntax, and semantics (Auger & Steele; Lipson; Mesthrie; Meyerhoff), and discussions of the formal linguistic underpinnings of regional dialect variants (Mesthrie), diachronic language states (Auger & Steele, Lipson), and learning stages (Lee). As the range of subjects suggests, variationist linguistics is not so much a discrete area of inquiry as a perspective usefully applied to many types of language-related research. This approach is a hallmark of the linguistics program at Penn, and for this reason the editors are grateful for the opportunity to uphold the emerging tradition of publishing NWAV(E) proceedings as a volume of the *Penn Working Papers in Linguistics*.

On behalf of the Working Papers committee, we would like to thank the eleven authors, the conference organizers and participants, and especially the departments in the linguistics program at the University of Georgia.

Christine Moisset and Mimi Lipson
Issue editors

Vowel Epenthesis in Vimeu Picard: A Preliminary Investigation[1]

Julie Auger and Jeffrey Steele

1 Introduction

One of the most striking phonological features of Picard, a Gallo-Romance language spoken in Northern France, is the apparent inversion of an unstressed vowel with the consonant that precedes it. This phenomenon is illustrated in (1) with several French words and phrases and their Picard equivalents in the Vimeu variety[2]:

(1)	French	Vimeu Picard	
a.	*grenouille*	*gu**é**rnouille*	'frog'
b.	*comme **des** harengs*	*conme **édz** hérins*	'like herrings'
c.	***Je** n'ai pas le temps*	***Éj** n'ai point l'temps*	'I don't have time'

While this phenomenon is attested in many varieties of colloquial French (e.g., Picard 1991 for Québécois, Poirier 1928 for Acadian, Lyche 1995 for Cajun, and Morin 1987 for Parisian) as well as in other Gallo-Romance dialects (e.g., Francard 1981 for Walloon and Spence 1990 for Norman), it is, to our knowledge, nowhere as common or regular as in Picard. Indeed, while metathesis is restricted to a few segments in other Gallo-Romance varieties, in Picard,

---

[1] The research for this paper was supported in part by an FCAR grant no. NC-1648, a research grant from the College of Arts and Sciences at Indiana University, and a Summer Faculty Fellowship from Indiana University to the first author. We would like to thank the following people for answering our numerous questions on epenthesis and syllable structure: Stuart Davis, Ken de Jong, Dan Dinnsen, Heather Goad, Glyne Piggott, and Yvan Rose. We would also like to thank Jean-Pierre Calais and Jean-Luc Vigneux for their help with the Picard data. Finally, we thank Rea Janezich, Angela Markley, and Kelly Sax for their help in gathering data and their efforts in finding relevant examples for testing our hypotheses. As usual, none of these people should be held responsible for any shortcomings in this paper.

[2] Vimeu is located in the westernmost area of the Somme department; it is delimited by the Somme river to the north, the Bresle river (and Normandy) to the south, the English Channel to the west, and departmental road 901 to the east.

U. Penn Working Papers in Linguistics, Volume 6.2, 1999

BEST COPY AVAILABLE

/e/ may appear before any consonant (e.g., /d/ in (1)b and /ʒ/ in (1)c); as we will see, this movement is obligatory in many contexts.

The present paper constitutes a preliminary investigation of the complex interplay of linguistic factors which govern this phenomenon in Vimeu Picard (VP). First, we will examine the data in (1) above in order to see whether they are the manifestation of the same phonological phenomenon or whether they result from different phonological processes. Our conclusion will be that underived word-internal epenthesis must be distinguished from the word-boundary and monosyllabic word phenomena. We will then focus on the word-boundary and monosyllabic-word data, arguing that the apparent reversal of consonant and vowel is best described as epenthesis rather than metathesis (cf. Picard 1991 and Lyche 1995 for similar analyses of *re-* in Québécois and in Cajun). Finally, we will examine the syllable structure of VP and show that epenthetic vowels are inserted when they are required in order to save consonants that could not otherwise be syllabified.

2 Word-internal vs. Boundary Inversion: One Phenomenon?

Many pairs of French-Picard words similar to the pair in (1)a can be found:

(2)		French	Vimeu Picard	Picard	Gloss Picard word
	a.	*grenier*	*guérnier*	[gernje]	'attic'
	b.	*brebis*	*bérbis*	[berbi]	'sheep'
	c.	*crêpe*	*quérpette*	[kerpɛt]	'pancake'

Knowing, for example, that the etymon for *grenier/guérnier* is the Latin *granarium*, it is tempting to posit a general metathesis rule which derives the Picard forms from a French-like underlying form /grenje/. However, this analysis does not account for all of the forms above. One problem is *bérbis*: in this case, it is the French word which has undergone metathesis, at least historically, since the etymon for this word is the Proto-Romance form **berbicem*. Moreover, certain Picard words exhibit variation between a metathesized and a non-metathesized form (e.g., *quérpette* is attested alongside *crépette* for 'pancake'), while other words are attested only in their apparently metathesized form. Finally, and most importantly, many Picard words containing a /Cre/ sequence cannot be metathesized, as shown in (3):

(3)			
	a.	*prépérer/*pérpérer*	'to prepare'
	b.	*adréche/*adérche*	'address'

While more research would be necessary to provide a definitive analysis of non-derived word-internal metathesis in VP, we propose that *guérnouille*, *guérnier*, and *quérpette* are all the result of a historical process of metathesis which intermittently inverted /re/ sequences. This analysis accounts for the fact that only /re/ is metathesized word-internally in Picard (never /le/, for instance), that not all /Cre/ sequences are metathesized, and that some words exhibit variation between /Cre/ and /Cer/. In addition, if we assume that this phonological change took place in the not-so-recent history of Gallo-Romance, we account for the fact that the words for 'attic' and 'frog' contain metathesized sequences in many Gallo-Romance varieties.

We now turn to the word-initial and monosyllabic-word environments. Contrary to the word-internal context, the apparent metathesis phenomenon is not limited to /r/ in this context, as shown in (4) and (5) below:

(4)		Picard	French cognate	Picard gloss
	a.	*ébzoin*	*besoin*	'need'
	b.	*érvénir*	*revenir*	'to come back'
	c.	*élveu*	*lever*	'to raise'
	d.	*évnir*	*venir*	'to come'
	e.	*écmin*	*chemin*	'road'
	f.	*émnaceu*	*menacer*	'to threaten'
(5)	a.	*vir él portrait*	*voir le portrait*	'to-see the picture'
	b.	*jours éd vagances*	*jours de vacances*	'days of vacation'
	c.	*aveuc és castchètte*	*avec sa casquette*	'with his cap'
	d.	*j'passe éch mércrédi*	*j'passe ce mercredi*	'I spend Wednesdays'
	e.	*Albért, émn honme*	*Albert, mon homme*	'Albert, my husband'
	f.	*pu fort éq li*	*plus fort que lui*	'stronger than him'

Once again, however, a synchronic metathesis rule fails to account for the data, since such a general phonological process would generate ungrammatical forms. (6) below contains pairs of words which differ only, or most importantly for us, in the position of /e/, thus showing that not all /e/'s can be metathesized.

(6) a. *dégouteu* 'to disgust'   c. *ménageu* 'to spare'
b. *édgoutteu* 'to drip'   d. *émnaceu* 'to threaten'

(7) a. *Et pi qu'chés guérnouilles i croassoait't* [**kʃeg**ernuj]/*[**keʃg**ernuj]
'and since the frogs cawed [sic]'

b. *six éch n'est point coér neuf* [siseʃne]/*[sisʃene]
'six it NEG is not still nine' = 'six is still not nine'

Considering these data, the possibility emerges that our analysis of word-internal metathesis above was too narrow in scope and that we should consider the words in (4) and (5) above as additional examples of historical metathesis. There are, however, two arguments against this analysis. First, while word-internal metathesis is common in Gallo-Romance varieties, metathesis in word-initial position and monosyllabic words seems to be restricted to Picard. Second, the "metathesized" vowel in words like *émnaceu* and *éch* is not always present, and its presence is predictable based on the environment: as shown in (8) below, /e/ appears after a consonant, but not after a vowel.

(8)	a.	*boéne* ***ésmaine***	[bwen esmɛn]	'good week'
		in ***smainne***	[ɛ̃ smɛn]	'during the week'
	b.	*l'frére d'**ém** feume*	[l frer d em fœm]	'my wife's brother'
		*vlo **m**'valise*	[vlo m valiz]	'here's my suitcase'

The question now becomes: which is the underlying form? I.e., is a vowel inserted following a consonant, or is a vowel deleted following another vowel? Both processes are widely attested. There is evidence that, in this case, the consonant-initial forms are underlying. Many grammatical words in Picard consist of two allomorphs: one occurring prevocalically, the other preconsonantally. E.g., the masc.sg form for the definite determiner is *chu* before a consonant and *chl'* before a vowel, while the 3masc.sg subject marker is *il* before a vowel and *i* before a consonant, as shown in (9)a and (9)b below. Thus, we can use this test to determine whether words like *cmin/écmin* 'road' and *cmincher/écmincher* 'to begin' have vowel- or consonant-initial underlying forms. (9)c shows that the preconsonantal allomorphs are selected.[3]

(9)	a.	*chu bal*	'the bal'	*chl'autocar*	'the bus'
	b.	*i court*	'he runs'	*il avoait*	'he had'
	c.	*chu cmin/*?chl'écmin*		'the road'	
		i cminche/?il écminche*		'he starts'	

Based on these facts, we conclude that the underlying form for *cmin* is /kmɛ̃/ and that a word-initial /e/ is inserted when *cmin* follows a consonant. Thus, the minimal pairs in (6) and (7) above can be attributed to the fact that some words have underlying /e/'s, while others acquire a vowel in certain phonological contexts; e.g., the underlying form for *dégouteu* is /degutø/, while that for *édgoutteu* is /dgutø/. Therefore, what looks, by comparison with French, like metathesis is actually a process of word-initial vowel epenthesis.

---

[3] Another piece of evidence against vowel deletion is the fact that the alternation in (8) is completely regular and thus differs from the related phenomenon of vowel aphaeresis, which affects all vowels and is variable in VP.

Now that we have established that we are witnessing epenthesis rather than metathesis, we should examine one final context where a vowel is inserted between consonants. Indeed, it is not rare to see an /e/ appear at the end of words when words ending in consonant clusters precede a consonant, as in (10) below. Given that this environment mirrors the word-initial context, we assume that they are both instances of the same general epenthesis rule.

(10) a. *quante al sorte* [sɔrt] *aveuc Sidonie*
'when she goes-out with Sidonie'
b. *J' én sorté* [sɔrte] *point*
'I NEG go-out not' = 'I'm not coming out'

3 The Syllable Structure of Vimeu Picard

In the preceding section we saw that epenthetic /e/ is inserted at the edges of consonant-initial and consonant-final words when these words surface in contact with other consonants. Thus, vowel epenthesis in VP appears to play the same role as in many other languages: it allows otherwise unsyllabifiable consonants to be syllabified (cf., e.g., Clements & Keyser 1983, Itô 1989, Repetti 1996, Colina 1997). In order to ensure that this analysis is warranted for Picard, we must, however, determine the syllable structure of this language. Assuming such widely accepted principles as the Sonority Sequencing Principle (Selkirk 1984), which stipulates that segments in a sequence are syllabified with the most sonorous segment constituting the head or nucleus of the syllable with a decrease in sonority as one moves towards either edge, the Sonority Hierarchy (e.g., Clements 1990), which ranks segments in relative increasing sonority from obstruents to nasals to liquids to glides to vowels, and Place of Articulation Restrictions (e.g., Rice 1992), which stipulate that segments syllabified within the same constituent may not bear the same place of articulation, we provide the following description of the syllable structure of VP.

(11) a. Branching onsets require a distance of 2 between their components

i.	/pl/	*m'plache*	[m.plaʃ]	'my place'
ii.	/pj/	*ch'piot*	[ʃ.pjo]	'the kid'
iii.	*/km/	*_ch' cmin_	*[ʃ.kmɛ̃]	'the road'
iv.	*/dv/	*_route dvant_	*[rut.dvɑ̃]	'road ahead'

b. Branching nuclei consisting of [w,ɥ] + vowel are allowed, but not those consisting of [j] + vowel

i.	_troés_	[trwe]	'three'
ii.	_pluie_	[plɥi]	'rain'
iii.	_oblieu_	*[obljø]/[obliø]	'to forget'

c. Complex codas are allowed, if the 1st element is a liquid[4]

i.	_pérc_	[perk]	'park'
ii.	_calme_	[kalm]	'calm'
iii.	*_présque/présqué_	[*presk/pres.ke]	'almost'

d. Codas may license place features

acceptabe	[ak.sɛp.tab]	'acceptable'

e. Appendices:

i. /t/ in word-final position: /t/ is the only consonant which can occur at the end of all words, regardless of the sonority of the preceding segment (compare with /ʃ/, the subjunctive morpheme, which cannot occur after an obstruent)

in dirécte	[ɛ̃.di.rekt]	'in direct' = 'live'
i conmuniq't	[i.kɔ.my.nikt]	'they communicate'
qu'j'ém dépéque/*dépéquche		'that I hurry.subj'
qué j'dorche		'that I sleep.subj'

ii. no word-initial appendix: contrary to French and English, word-initial appendix /s/ is not allowed; instead, this consonant behaves like other consonants and requires an epenthetic vowel as syllabic support

*_inne statue/inne éstatue_	[*ɛ̃n.sta.ty/ɛ̃.nes.ta.ty]	'a statue'

f. Syllables with empty nuclei are not allowed: either consonant deletion or vowel epenthesis is required

*_contre/conte/contré_	*[kɔ̃tr] / [kɔ̃t] / [kɔ̃.tre]	'against'
pour *cmincher/écmincher	*[pur.kmɛ̃.ʃe]/[pu.rek.mɛ̃.ʃe]	'to start'

The discussion and the examples above show how syllable structure and epenthesis are closely intertwined; indeed, it is often the fact that epenthesis is required in some contexts but not in others that reveals differences in syllabification. E.g., the fact that /sr/ is a possible complex onset[5] but that /sm/ is not is revealed by the contrastive behavior of the two sequences when they follow a consonant: *inne sringue* [ɛ̃n.srɛ̃g] 'a syringe' vs. *inne ésmainne* [ɛ̃.nes.mɛ̃n] 'a week'. While this approach to determining syllable structure might appear circular, the fact that the structure which it reveals falls so clearly in line with general principles of syllable structure makes us confident that our conclusions are valid. Indeed, these conclusions allow us to account for a number of facts which had initially appeared mysterious to us. E.g., based on the fact that there are no /vl/ onsets in French, we were surprised to discover that no epenthetic vowel is inserted when *vlo* 'here's' follows a consonant. The syllable structure above accounts for this: /vl/ is a possible onset in VP[6], making epenthesis unnecessary. Conversely, with the word *rién* 'nothing', we were surprised to find that epenthesis is obligatory when this word follows a consonant-final word. The requirement for a minimal distance in the onset predicts obligatory epenthesis, since, according to Clements' Sonority Hierarchy, /rj/ is not a possible onset. The contrast between **ch' rio* 'the stream' (the grammatical form is *chu rio*) and *ch' piot* 'the kid' confirms that /rj/ and /pj/ are syllabified differently: /pj/ is a possible onset in VP, while /rj/ is not.

---

[4] Branching codas are generally not accepted in phonological theory; for example, Kaye's 1990 Binary Principle forces all such word-final consonants to be syllabified as onsets.

4 Vowel Epenthesis and Syllable Structure

Our hypothesis concerning vowel epenthesis in VP is the following: a vowel is inserted in order to syllabify a segment which would otherwise be unsyllabifiable for reasons of sonority or place of articulation. While some languages delete unsyllabifiable consonants, it would seem that this option is generally disfavored in Picard. In the current terms of Optimality Theory, it would appear that VP ranks MAX, the constraint which requires every element of the input to be overtly realized, higher than DEP, the constraint which requires every element of the output to have a correspondent in the input.

In this section, we will test our hypothesis against the epenthesis facts of VP. We will make use of the syllable structure sketched in (11) above, relying particularly on the observation that this language does not normally allow syllables with empty nuclei. We will also make use of the notion of Intonational Phrase (IntPhr) as defined by Selkirk 1995:566 : IntPhr's are "spans of the utterance which are delimited by boundary tones". For example, while there is a single IntPhr in _I called Mary yesterday_, there are normally two such phrases when a sentence contains a preposed adverbial complement (_Yesterday, I called Mary_) or a left-dislocated phrase (_Mary, I called her yesterday_). As we will see, IntPhr boundaries play a crucial role in determining whether epenthesis is obligatory or variable.

---

[5] While, phonetically, /r/ is either apical or uvular in VP, we consider its phonological behavior to be that of a placeless glide, following Rice 1992.

[6] The question thus becomes why this onset is not possible in English or French.

The simplest case of vowel epenthesis is found in word-initial position within an IntPhr: in those cases where a consonant cannot be syllabified within the onset of the word to which it belongs underlyingly because of place and/or sonority restrictions, an epenthetic vowel is inserted allowing the consonant in question to be syllabified as the coda of a new syllable. This is exemplified in (12):

(12) a. _in liméro comme ézz/*zz eutes_ [kɔ.mez.zøt]
'an issue like the others' = 'an issue like others'
b. _pour écmincher/*cmincher_ [pu.rek.mɛ̃.ʃe] 'to start'

When a word beginning with a consonant that cannot be syllabified as part of the onset is located at the beginning of an IntPhr, vowel epenthesis becomes optional. As shown in (13) below, this variable behavior is found in sentence-initial position as well as after an intonational break within a sentence. Our hypothesis based on syllable structure does not predict that there should be a difference between these and the word-initial cases in (12) above, since, in both contexts, a consonant lacks a syllabic nucleus that can support it. We would like to propose here that the phrase-initial consonant can be licensed either directly by the syllable created by the insertion of an epenthetic vowel or indirectly by the IntPhr itself. While admitting that this proposal is in need of further development, it is not uncommon for segments to be licensed indirectly at the edge of a prosodic domain, whether that domain be the foot, the prosodic word, or higher up in the prosodic hierarchy (Piggott 1998).

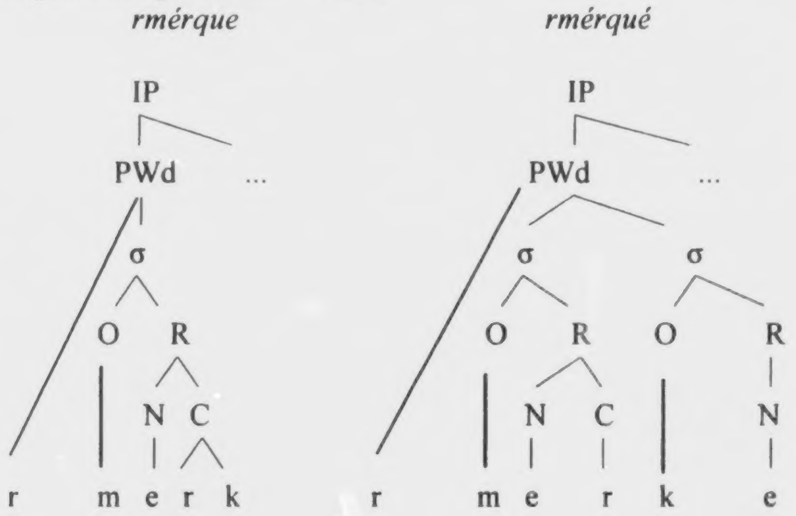
(13) a. _Ch'/Éch troisième honme il a parti_ 'The third man has left'
b. _Ben non, éj/j' n'ai point peu_ 'Eh no, I couldn't'

In word-final position, the situation is more complex due to the fact that different syllabification options exist. Where only one strategy is available, a single output is possible. E.g., when the final consonants do not constitute a possible coda, and when the second consonant is not /t/ and thus not a possible appendix, only an epenthetic vowel can save the final consonant. This is the case, for instance, when the first of the two consonants is not a liquid and the second is not /t/, the only possible appendix, as shown in (14):

(14) a. *pasqué/*pasque j'sus pu souvint din chés cambes*
'because I'm more often in the bedrooms'
b. *O n'ouéyouot présqué/*presque pu clair*
'we NEG saw almost anymore clear'
= 'we could barely see anything anymore'

When the final cluster is a possible word-final consonant cluster, two outputs are also possible. This situation arises in two different cases. One is when the first consonant of the sequence is a liquid, since this is the only type of coda cluster that is possible in VP. In such cases, the final consonant can be syllabified as part of the coda or as the onset of another syllable if an epenthetic vowel is inserted, as shown in (15).

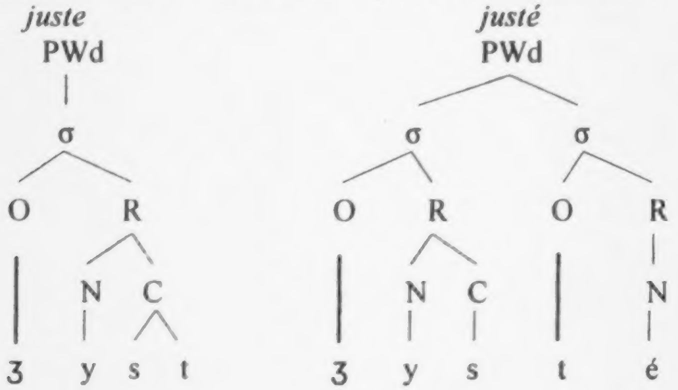
(15) *Rmérque/Rmérqué bién...* 'Note...'


rmérque *rmérqué*

Yet why should there be two outputs? We propose that this results from the fact that neither output is truly optimal, each of them violating a highly-ranked constraint: *rmérque* violates Kaye's 1990 Binary condition which prohibits complex codas, while *rmérqué* violates DEP, since the epenthetic vowel lacks a correspondent in the input form. In future work, we intend to explore the possibility that this variation could be captured through the notion of crucial unranking (Prince & Smolensky 1993), following Anttila's 1997 proposal.

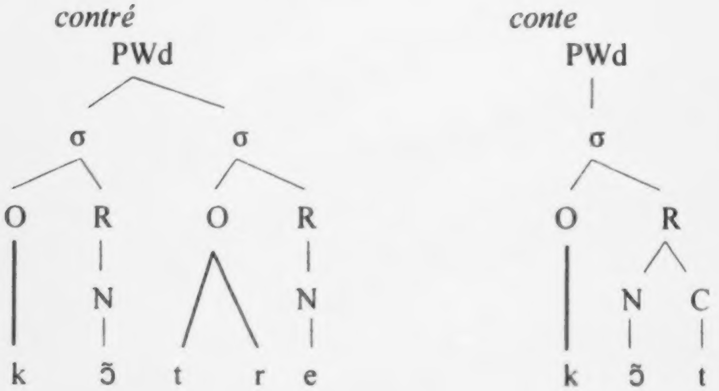
When the final consonant is /t/ and the preceding consonant is not a liquid, as in *juste* 'just', the cluster is not a possible coda; however, /t/ can be licensed by the prosodic word (PWd) as an appendix. In this case also, two outputs are possible: when the following segment is a consonant, *juste* can either surface without an /e/, if the /t/ is licensed by the PWd as an appendix, or with an /e/, as shown in (16). Once again, we hypothesize that the variation between the two outputs is due to the fact that they violate constraints which are not crucially ranked with respect to each other: as indirect licensing is marked vis-à-vis direct licensing, the *juste* form, which contains an appendix indirectly licensed by the PWd, is not optimal, while the epenthetic vowel in *justé* violates the DEP constraint.

(16) *juste/justé derrière éch chauffeur* 'just behind the driver'



One final context where word-final vowel epenthesis is possible within an IntPhr involves consonant clusters that have an onset-like profile rather than a coda-like profile, as in, e.g., *contre* /kɔ̃tr/ 'against'. /tr/ is not a possible coda, because of its sonority profile, and /r/ is not a possible appendix in Picard. Two output forms are possible. First, as in the previous cases, an epenthetic vowel can be inserted making it possible for /tr/ to be syllabified as an onset cluster, as shown in the *contré* form in (17) below. It is also possible, however, to delete the final liquid and to syllabify the /t/ as part of the coda of the preceding syllable, thus yielding a monosyllabic output form: [kɔ̃t].

(17)



The latter form, *conte*, raises a question: why is it possible to delete the /r/? Deletion was never a possibility for any of the other contexts considered so far. In our data, word-initial consonants are never deleted; some final consonants can be. However, the liquids in words like *contre* and *sable* 'sand' are the only segments which seem to be deleted in synchrony. Other consonants, such as the /m/ that might be posited at the end of *catéchisse* 'catechism' or the /t/ at the end of *journalisse* 'journalist', do not appear to be present in the underlying forms of these words. Indeed, since it is possible to insert epenthetic vowels to save final consonants, we should find instances of *catéchismé* and *journalisté* in our data. No such examples can be found. On the other hand, examples of *contré* and *sablé* were present, showing that the final liquids are present in the underlying forms of these words. Space constraints prevent us from investigating this issue in detail here. However, we attribute the fact that only word-final consonants can be deleted to the word-recognition difficulties that would arise if the initial consonant of a word were missing (e.g., [mɛ̃] for *cmin*). What remains to be determined is the reason why deletion of a final consonant is possible in *contre* but not in *pasque* (*[pas]).

Before moving on to the last context, that of word-final clusters occurring at the end of an IntPhr, we would like to briefly discuss an apparently surprising environment: that of 3pl verb forms. We mentioned earlier that the 3pl morpheme is a /t/ segment that can be added to any type of stem, no matter whether it ends in a vowel or in a consonant. When the stem is consonant-final, an epenthetic vowel is variably inserted if the verb occurs before a consonant-initial word, as seen in (18). In this respect, 3pl verb forms behave like any underived word which ends in a possible word-final consonant cluster.

(18) a. *ses amis qu'i voz* ***apport't*** *vo jornal*
'his friends that to-you bring your newspaper'
= 'his friends who bring you your newspaper'
b. *pi i* ***sort 'té*** *touté deux*
'and they go-out all two' = 'and both of them leave'

Surprisingly, however, an epenthetic vowel can be inserted even when the 3pl morpheme is not preceded by a consonant, as in (19):

(19) a. *il avoait'té voté* 'they had voted'
b. *pér nuit il avoait't voté* 'at night they had voted'

This is the only context where epenthetic /e/ regularly appears despite the fact that it is not required to break up sequences of consonants unpronounceable in

Picard. Why might this be? The presence of the epenthetic vowel cannot be attributed to the fact that /t/ is an inflectional morpheme, since the 2pl /t/ morpheme does not trigger epenthesis, as we can see in (20):

(20) a. *O sroéte putot 'conte' ?* 'you.pl would.be.2pl rather against'
b. *o n'comprindroéte point* 'you.pl would.understand.2pl not'

Rather, we must take the spelling which is found in (19) seriously: the *t't* reflects the geminate character of the 3pl ending, which should, in consequence, be /tt/ rather than simply /t/. VP contains a number of morphemes which consist of a geminate consonant; for instance, the 3sg.acc pronominal clitic is /ll/, and the partitive/genitive pronominal clitic is /nn/. Both of these forms share the same distribution: while the geminate form surfaces in intervocalic environments, as in (21)a below, a simplified form is found in contact with a consonant, as shown in (21)b. However, degemination is not obligatory in this context: if an /e/ is inserted between the geminate and the consonant, the geminate is preserved, as in (21)c.

(21) a. *O ll'avoéme attindu* 'we it had waited'
b. *tu l'sais* 'you.sg it know'
c. *I llé savoait* 'he it knew'

Thus, there is evidence for a general but variable degemination rule in VP. Furthermore, Éloy 1997 reports that, in the Picard of Amiens, 3pl /tt/ is variably degeminated. Consequently, we propose that variable epenthesis in forms like *avoait't* correlates with variable degemination: a vowel is inserted when a geminate /tt/ precedes a consonant, but not when the geminate is simplified. In contrast, the morpheme for 2pl is an ungeminated /t/ which does not require epenthesis.

There is one final environment where we would expect epenthesis to occur: in word-final position at the end of an IntPhr. Very interestingly, it appears that this environment is generally incompatible with epenthetic /e/. As we can see in (22), words which normally allow for epenthesis within an IntPhr must surface without an epenthetic vowel at the end of an IntPhr.

(22) a. *I povoait rpinser à s'nouvelle rinconte/*rincontre/*rincontré*
'He could think back to his new encounter'
b. *din chés leumiéres qu'i trann't/*trann'té*
'in the lights that flicker.3pl'

While we can only conjecture on the reason for this prohibition at this point, one plausible explanation is that epenthetic vowels cannot be stressed in Picard. An alternative explanation would invoke an ALIGN RIGHT constraint which would force the end of a morphological word to be aligned with the end of the IntPhr (McCarthy & Prince 1993).

One final question remains: why do postverbal clitics constitute an exception, as in (23)? The fact that 3pl verbs do not accept an epenthetic /e/ in this context rules out the possibility that we might be dealing with the same phenomenon as in (19) and (21) above. Another possibility is that clitics do not constitute a real exception. Indeed, it is common for pronominal clitics in Gallo-Romance languages to have different preverbal and postverbal forms. E.g., in French, *me* is preverbal and *moi* is postverbal for 'me'. In VP, two forms must be distinguished, too: the preverbal form is a single /m/, while the postverbal form is a geminate /mm/. Consequently, it is not out of the question that the input form for 1sg.acc/dat in postverbal position could contain an /e/ that would be underlying rather than epenthetic: /mme/. Indeed, we must not forget that not all /e/'s in Picard are epenthetic and that underlying /e/'s are allowed in the final position of an IntPhr, as shown in (24).

(23) *Mais diseu mmé, quoé qu'i foait lo*
'but tell me, what is he doing there?'
(24) *Il ont mingè des quérpèttes au solé.*
'they have eaten of-the pancakes at-the sun'
= 'They ate pancakes in the sun'

5 Conclusion

We have shown in this paper that what at first may have appeared to be metathesis consists of two different phenomena: historical word-internal metathesis and synchronic vowel epenthesis. We have focused on the latter and shown that, as with cases of epenthesis in general, a segment, here /e/, is inserted in order to syllabifiy consonants occurring at edges of domains that would be otherwise unsyllabifiable and thus subject to stray erasure. We have demonstrated that epenthesis is generally predictable once the syllabic structure of Picard has been determined. We have also identified contexts which allow for variable epenthesis. In these cases, other phonological properties of Picard, including the ability of the PWd and IntPhr to act as indirect licensers, allow for variable outputs.

This analysis has raised a number of questions which require further investigation. As we expand the object of our inquiry to include clitic sequences and derived words, we expect to answer these questions and refine our analysis in general. Through a meticulous study of the variable patterns, we also expect to contribute to a very interesting research trend which proposes a model of linguistic competence capable of generating language-internal variation (cf., e.g., Anttila 1997; Sells, Rickford, & Wasow 1996).

References

Anttila, Arto. 1997. Deriving variation from grammar: A study of Finnish genitives. In *Variation, Change, and phonological theory*, ed. F. Hinskens et al., 35-68, Amsterdam: Benjamins.

Clements, George N. 1990. The role of the sonority cycle in core syllabification. In *Papers in laboratory phonology I*, ed. J. Kingston & M. E. Beckman, 283-333. Cambridge: Cambridge University Press.

Clements, George N. & Samuel Jay Keyser. 1983. *CV phonology*. Cambridge, MA: MIT Press.

Colina, Sonia. 1997. Epenthesis and deletion in Galician: an optimality-theoretic approach. In *Issues in the phonology and morphology of the major Iberian languages*, ed. G. F. Martinez & F. A. Morales, 235-267. Washington, D.C.: Georgetown University Press.

Éloy, Jean-Michel. 1997. *La constitution du picard: une approche de la notion de langue*. Louvain-la-Neuve: Peeters.

Francard, Michel. 1981. Voyelles instables en wallon: propositions pour une approche globale. In *Dialectologie en Wallonie*, ed. Willy Bal, 169-200. Louvain-la-Neuve: Cabay.

Itô, Junko. 1986. Syllable theory in Prosodic Phonology, Doctoral dissertation, University of Massachusetts, Amherst.

Kaye, Jonathan. 1990. *Coda* licensing. *Phonology* 7:301-330.

Kenstowicz, Michael. 1993. Syllabification in Chuckchee: A constraints-based approach. *Formal Linguistics Society of Midamerica* IV, 160-181.

Lyche, Chantal. 1995. Schwa metathesis in Cajun French. *Folia Linguistica* 29:369-393.

McCarthy, John & Alan Prince. 1993. *Prosodic Morphology I*. Technical Report #3, Rutgers University Center for Cognitive Science.

Morin, Yves-Charles. 1987. French data and phonological theory. *Linguistics* 25:815-843.

Murray, Robert & Theo Vennemann. 1983. Sound change and syllable structure in Germanic phonology. *Language* 59:514-528.

Picard, Marc. 1991. La *loi des trois consonnes* et la chute du cheva en québécois. *Revue québécoise de linguistique* 20,2:35-48.

Piggott, Glyne L. 1998. "At the right edge of words". Ms. McGill University.

Poirier, Pascal. 1928. *Le parler franco-acadien et ses origines*. Québec: Imprimerie franciscaine missionnaire.

Prince, Alan & Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. Ms. Rutgers University & University of Colorado, Boulder.

Repetti, Lori. 1996. Syllabification and unsyllabified consonants in Emilian and Romagnol dialects. In *Aspects of Romance linguistics*, ed. C. Parodi et al. 373-382. Washington, D.C.: Georgetown University Press.

Rice, Karen. 1992. On deriving sonority: A structural account of sonority relationships. *Phonology* 9:61-99.

Rose, Sharon. 1997. Theoretical issues in comparative Ethio-Semitic phonology and morphology. Doctoral. dissertation, McGill University.

Selkirk, Elisabeth. 1984. On the major class features and syllable theory. In *Language sound structure*, ed. M. Aronoff & R. T. Oehrle, 107-136. Cambridge, MA: MIT Press.

Selkirk, Elisabeth. 1995. Sentence prosody: Intonation, stress, and phrasing. In *The handbook of phonological theory*, ed. J. A. Goldsmith, 550-569. Cambridge, MA: Blackwell.

Sells, Peter, John Rickford, & Thomas Wasow. 1996. An Optimality Theoretic approach to variation in negative inversion in AAVE. Natural Language and Linguistic Theory 14:591-627.

Spence, Nicol. 1990. Sporadic changes in Jersey French. In *Variation and change in French*, eds. J. N. Green & W. Ayres-Bennett, 210-225. London: Routledge.

Julie Auger
Dept. of French & Italian
Indiana University
Bloomington, IN 47405
jauger@indiana.edu

Jeffrey Steele
Dept. of Linguistics
McGill University
Montréal, Québec
Canada, H3A 1G5
jsteel1@po-box.mcgill.ca

The *University of Pennsylvania Working Papers in Linguistics (PWPL)* is an occasional series produced by the Penn Linguistics Club, the graduate student organization of the Linguistics Department of the University of Pennsylvania.

Publication in this volume does not preclude submission of papers elsewhere; copyright is retained by the authors of the individual papers.

Volumes of the Working Papers are available for $15, prepaid. Please see our web site for additional information.

The PWPL Series Editors

Jim Alexander
Alexis Dimitriadis
Na-Rae Han
Elsi Kaiser
Michelle Minnick Fox
Christine Moisset
Alexander Williams

How to reach the PWPL

U. Penn Working Papers in Linguistics
Department of Linguistics
619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104–6305

http://www.ling.upenn.edu/papers/pwpl.html
working-papers@ling.upenn.edu

Lexical Borrowings from French in Written Quebec English: Perspectives on Motivation

Pamela Grant-Russell and Céline Beaudet

The purpose of this paper is to explore perspectives on motivation for lexical borrowings from French in written Quebec English. To do so, we propose first to examine factors that account for the presence and patterns of loanwords in Quebec English in general and then to identify factors that motivate the use of loanwords as code choices in individual texts.

1 Background

On the federal level, Canada is officially bilingual, the Official Languages Act of 1969 having established English and French as the two official languages of Parliament and federal government services. However, French has been and remains the language of the minority of Canadians, other than in Quebec, the one predominantly French province. In Quebec, where French is the first language of more than 80% of the population, French is the only official language of the province.

As a minority language not only within Canada but also within the whole of North America, French has for centuries been on the defensive against assimilation by English. In the last thirty-odd years, the linguistic balance of power has changed radically within the province of Quebec. The Quiet Revolution, the modernization of Quebec society in the 1960s, led to the rise of the Québécois nationalist movement based largely on the protection and promotion of French within Quebec as a means of ensuring the survival of this French-speaking community in North America. The early 1960s saw the beginning of language planning in Quebec, with the creation of the *Office de la langue française*. In 1974, the *Loi sur la langue officielle* made French the only official language of Quebec. The dominant position of French was firmly established by Bill 101, the *Charte de la langue française*, in 1977. Since then, Quebec has seen concentrated efforts and aggressive language policies to ensure the use of French in Quebec in all walks of life. On the academic level, the interest in Quebec French and the influence of English on it has been enormous. Bourhis states, "Quebec...remains the most active centre for the investigation of language attitudes in the Francophone world" (1982:35). Recent years have seen the publication of literally thousands of articles and monographs on Quebec French, as well as various dictionaries on Quebec usage (Poirier 1985; Beauchemin, Martel and Théoret 1992).

Interest in Quebec English has been much more modest and relatively recent. Although the borrowing of French-Canadian words into Canadian English had been observed since colonial times (Avis 1978), earlier studies of English in Quebec (such as Hamilton 1958 and sections of Orkin 1971) focused mainly on the influence of British and American Englishes. In the past two decades, scholarly interest in Quebec English has grown. Articles on the topic include papers by Roberts (1982), Manning and Eatock (1982), Plaice (1984), Palmer and Harris (1990), Fee (1991, 1995), Russell (1996) and Grant-Russell (1998), as well as a monograph (McArthur 1989). Recently published reference works on the English language have acknowledged the uniqueness of Quebec English (McArthur 1992; Fee and McAlpine 1997). Less scholarly commentary on the phenomenon can be found in countless articles and editorials in the Quebec English press, as well as in publications such as *Bonjour Quebec eh?* (Keith-Ryan and McCully 1996) and the *Angloman* comic books (Shainblum and Morrissette 1995, 1996).

2 Description of Study

These findings emerge from a research project underway at the Université de Sherbrooke,[1] based on the analysis of lexical borrowings from French in a written corpus of English-language journalistic, administrative, touristic, governmental, and other non-literary texts. The findings are based on a partial analysis of an expanding corpus in an ongoing research project, and our analysis and conclusions are obviously not definitive.

The corpus analysed to date consists of three components: a journalistic corpus, a touristic corpus, and an administrative corpus. The journalistic corpus is by far the most extensive, and is composed of both electronic and paper documents. The electronic corpus analysed to date contains the 1994 and 1995 issues of *The Gazette*, the principal English-language daily of Quebec, containing 120,447 newspaper articles. We have been searching this section of the corpus electronically to ascertain frequencies of uses of loanwords, to find examples of usages, and to compile quantitative data. This electronic corpus is supplemented by paper copies of a dozen weekly and monthly English-language newspapers from various areas throughout Quebec. The touristic corpus consists of over two hundred tourist guides, brochures, and pamphlets put out by government, municipalities, and tourist bureaus and agencies. The administrative corpus consists of some fifty documents, including annual reports and other publications, of English-language hospitals, English-language school boards, and English-language associations from throughout the province. Although we cannot provide a profile of the individual writers of these texts, we assume that they are literate, fluent speakers of English from Quebec.

---

[1] The authors gratefully acknowledge financial support for this research in the form of grants from *FCAR* (*Fonds pour la Formation de Chercheurs et l'Aide à la Recherche*) and the *Fondation de l'Université de Sherbrooke*. Thanks also to research assistants Elspeth Tullock, Marielle Mencé, Lucie Dubois and Nancy Wright.

Most studies of language variation have focused on spoken discourse; indeed, the dialectologists' preferred collection method is the recording of spontaneous speech, and the ideal corpus for studying variation has traditionally been the spoken vernacular, in which the minimum attention is paid to monitoring (Labov 1972:308). Thus, although the patterns and choices of lexical variants observed in our written corpus share many characteristics with those that manifest in speech, our use of a written corpus for language contact research has several important implications. First, we can expect to find a lower frequency of borrowing in our corpus than researchers would find in informal spoken discourse: informal speech reflects more interference than more formal discourse (Poplack 1983:113). Second, we can assume that the established borrowings used in the edited, published texts of our corpus have acquired a certain degree of acceptability, standardization, and authority: "Variants which are used in written communication are much more likely to be standardized than those which are used solely in the oral modality" (Bouchard Ryan, Giles and Sebastien 1982:3). Thirdly, since written discourse is often prepared with rhetorical forethought, we assume that some of the uses of borrowings reflect rhetorical intent. Fourthly, since our corpus reflects mediated discourse from the public domain, the writer is addressing an audience that is a synthesis of a real, heterogeneous group of readers and a fictitious readership invoked by the writer (Ede and Lunsford 1984 :167).

3 Findings to Date

This paper will focus primarily on full lexical borrowings, including single words such as *dépanneur* (convenience store), acronyms and initialisms such as *cegep* (a type of junior college) and *CLSC* (a local community service centre), and collocations such as *vieille souche* (a term used to designate French Quebecers of old stock).

Because of the great difference in size of the sections of our corpus, we established separate frequency lists for each section. We have observed that the semantic fields of borrowings are highly dependent on the section of corpus as well as on the genre of the text and the subject matter of the text in which they occur.

In the journalistic corpus, we recorded 45 high-frequency borrowed forms, which we defined, admittedly arbitrarily, as those occurring more than 10 times in this section of our corpus. Over one third (35%) of such words were from the fields of language and politics and designated language and ethnic group membership or political strategies (e.g. *allophone, francize, virage, nous, anglais,* and *pure laine*). Another 16% related to government, education, health, and social institutions and services (e.g. *cegep, régie*): these two semantic fields accounted for more than half of the established borrowings. Words from the fields of entertainment and tourism (e.g. *vernissage, auberge*) accounted for 20%; food (e.g. *poutine*), for 9%; transportation and daily life (e.g., *metro, autoroute, dépanneur*), for 11%; and 9% miscellaneous. The heavy concentration of terms related to language and politics is not surprising in view of the subject matter of a journalistic corpus, containing both news stories and editorials.

In the touristic section of the corpus, almost 90% of high-frequency borrowed forms (which we defined in this much smaller corpus as those having occurred more than 4 times) were concentrated in fields directly related to tourism: food, accommodation and dining (e.g. *gîte, auberge; terrasse*), history (e.g. *intendants, seigneur*), entertainment (e.g. *planche à voile, boîte à chansons*), and terms relating to natural features, fauna, flora, camping, fishing (e.g. *ouinaniche*), and hunting (e.g. *appelant, pourvoirie*). In the administrative section of the corpus, almost all borrowings were references to Quebec government departments, agencies, programs and services (eg. *francization, classes d'accueil, plan d'action*) or references to the educational and health care systems (e.g. *cegep, CLSC, régie*).

Thus what emerges is a profile of the domain of discourse as well as of the field of borrowing. Some established borrowings were found in all sections of the corpus (*cegep, régie, autoroute, metro)*; others were domain-specific, with high frequency use in specific contexts. For example, several of the most common journalistic borrowings from the field of language and politics (*pure laine, vieille souche*) were entirely absent from the administrative and touristic sections of the corpus. The occurrence of borrowings reflected not only the domain of discourse but the rhetorical or ideological dynamics of the individual text; for example, editorial-style articles and promotional writing contained higher levels of rhetorically motivated marked borrowings than did the administrative texts.

4 Motivation: An Analysis on Two Levels

In the following pages we will first examine factors that account for the presence and patterns of loanwords in the language of the Quebec English community, and we will then describe factors that motivate the use of loanwords as discursive choices in individual texts. Lexical borrowings can be perceived and described as both part of the language code shared by a community and as components of individual discourse, of language in use. This is the classic distinction of langue and parole. Grosjean (1982) differentiates speech borrowing and language borrowing; "Speech borrowing is what an individual does; language borrowing is done by a community" (Palmer and Harris 80). Borrowings that are assimilated into the language code are generally studied within the framework of historical linguistics and sociolinguistics as elements of language shift and language change. Fully assimilated borrowings become part of the community's repertoire and their presence in discourse may be unmarked and accounted for largely by factors external to the individual speaker.

On the level of discourse, the individual's choice of code from among the linguistic alternatives available in his or her repertoire can be seen as strategic, as indicative of the communicative intentions of the writer, the writer's attitudes toward the language communities in contact, and the writer's role vis-à-vis the audience (Myers-Scotton 1998; Benveniste 1966; Gérard-Naëf 1987). In this perspective, borrowings can reflect a situational, rhetorical or ideological dynamic and can be a manifestation of the writer's social identity or discursive intent. This is especially the case when borrowings have connotative and associative meaning or are exploited for rhetorical effect. Such considerations are addressed within various theoretical frameworks, including the linguistics of enunciation (Benveniste 1966) and genre theory (Swales 1990).

5 Large-Scale Motivations for Language Change

The presence of borrowings from French in Quebec English can be seen as the result of several large-scale factors: increased bilingualism in the English-language community; casual attitudes towards gallicisms among English-speakers; unilingual French designations for many governmental and institutional realities; the increased prestige of French; and lexical need and efficiency. These factors shape the repertoire of the common language of the Quebec English community.

5.1 Increased Bilingualism in the Host Language Community

With French being the language of the provincial government and the workplace, with the English-language school system promoting bilingualism at all age levels, and with the exodus of many of the anglophones who were unwilling to adapt to life in an officially French-speaking society, the level of bilingualism among Quebec anglophones has increased dramatically. According to Statistics Canada, English-French bilingualism among anglophones in Quebec increased from 36.7% in 1971 to 58.4% in 1991 to 61.7% in 1996. With more and more anglophones having two languages in their repertoires, borrowings from French are introduced more frequently and are spread and assimilated more rapidly throughout the community.

5.2 Attitudes Towards Borrowings

It is dangerous to generalize about Quebec anglophones' attitudes towards borrowings from French; as Crystal writes, "In an area where language attitudes are strong,...the extent to which Anglophone people are prepared to use words which are perceived to be French in origin varies greatly" (1995:343). Yet we can observe that, in contrast to the strong resistance to anglicisms in Quebec French, which Chantal Bouchard (1989) refers to as "une obsession nationale," gallicisms in Quebec English are seen as relatively benign. Although anglophones do have grave concerns about their future and their rights as a community in Quebec, the quality of the English language is not perceived to be at risk, with English-speakers having easy access to American and English-Canadian media, publications, and entertainment to support their use of standard English. Familiarity with the French language and culture are increasingly elements of the social identity of Anglo-Quebecers (see Reid 1998 for comments on the construction of Anglo-Québécoise identity in literature.) Although Roberts (1982) warns that the tendency towards Frenglish within the English-speaking Quebec community may be detrimental to that community's ability to communicate outside the province, the occurrence of gallicisms, particularly full lexical borrowings, has not been the subject of alarm. Interference is viewed as problematic primarily for learners of English as a second language, where it reflects a lack of competence in English. This nonchalance reflects the relatively casual attitude of English-speaking Canadians towards their language in general. In Orkin's words:

> [The] lack of concern which most English-Canadians show towards their daily speech contrasts sharply with the interest of a great many French-speaking Canadians in their distinctive variety of French.... French is the language of the minority only, with the result that the study of Canadian French has often been as much a matter of ethnic and political self-assertion as it has been of linguistic research. (1971:3)

5.3 Unilingual Institutional Designations

Most governmental, institutional, social and political entities in Quebec have unilingual French designations, creating a situation in which the referential reality is expressed only in French. All sections of our corpus contain a heavy concentration of French proper names of government departments and agencies, businesses, organizations, positions, and events. Many of these display standard French head-first word order (*Fête nationale, caisse populaire*). Also frequently observed is the use of acronyms (*cegep, ZEC, DEC*) and initialisms (*CLSC, SQ, PQ, BQ, OLF*) derived from French-language designations. The regular use of terms such as *régie, caisse,* and *fonds* in proper names has led to these terms being used generically in English as common nouns.

(1) ...they would pay for repairs only if the municipalities created a *régie* to manage the lake. (from *The Stanstead Journal*, 17/9/97, 8)

5.4 Increased Prestige of French

It has long been accepted that languages in general borrow from a language of higher socio-economic and political prestige. "The greater influence of one language over another can be accounted for by presuming a 'prestige differential' (Haugen 1969) that reflects the status of one language group vis-à-vis another" (Palmer and Harris 1990:77). English has increasingly assumed the status of a minority language within the province, leaving it more susceptible to borrowings from French. Quebec language policies have resulted not only in increased vitality, recognition and use of French as compared to English, but also in Quebec French being increasingly recognized as a legitimate and standard variety of French in its own right. Richard Bourhis observes, "The general atmosphere in favour of French in Quebec seems to have raised the status of Québécois French relative to both the English language and standard French" (1982:59).

The semantic fields of borrowings can be indicative of social and power relations between the communities in which they occur. In a study of borrowings from Arabic in the Algerian French press, Morsly states,

> L'emprunt, en situant quels objets, quelles valeurs sont adoptés ou non, deviendrait un indicateur des zones d'interculturalité et des zones de résistance. (1995:45)

The vocabulary of much of Quebec politics and linguistic debate has originated in French. Thus concentrations of borrowings are found in the fields of language and politics in Quebec English. Fee observes:

> The set of words that has been integrated most thoroughly into Quebec English and even beyond into Canadian and world English is the set of words that deals with Quebec politics, especially linguistic politics. Because much of the debate over these issues has been carried out in the national media and by some of the most important public figures in the country, the words used are quickly disseminated and integrated into the domain of Canadian political discourse. (1991:17)

5.5 Lexical Need

Lexical need and lexical efficiency have been suggested as motivations for some borrowing (Weinreich 1953). Terms which fill gaps in the host language lexicon by naming new objects or concepts can be distinguished from those naming objects or concepts for which the recipient language already has words; Myers-Scotton (1997:5) distinguishes these two types of borrowings by the terms *cultural borrowed lexemes* and *core borrowed lexemes*, the latter which she calls largely redundant. The borrowed word may express a cultural reality which has no equivalent in the host language community. In Quebec English, this is the case in particular with borrowings referring to Quebec history (*seigneur, intendant*) and politics (*étapisme; la grande noirceur; le beau risque; caribou*)

(2) Bloc leader Bouchard called, in April, for a "virage" or sharp turn in sovereignist strategy. (from Paul Wells, *The Gazette*, 10/6/95, B1)

However, not all borrowings motivated by lexical need are cultural borrowings; in some cases, a lexical gap does not indicate that a concept is culturally novel but simply that one language has a word for a specific concept where the other language doesn't. Vinay and Darbelnet examine various cases of and reasons for such gaps or lacunae as they relate to translation between French and English. "The signified may not exist or may not be acknowledged in one of the two languages; or it may exist in both but is only named independently in one of them" (1995 :65).

(3) Next week begins what the French call "la rentrée," the return to work, school – and reality. (from Nick Auf Der Maur, *The Gazette*, 30/8/95, A2)

6 Motivation for Borrowing in Discourse

A discourse contains various marks or indices, including lexical choices, that reflect the presence of the writer. Code choices that mark the discursive fabric of a text include switches of register, shifts of style, use of regionalisms, use of jargons from other domains, as well as interlingual borrowings. A full lexical borrowing, particularly one that is not presumed to be part of the audience's repertoire, can carry more than simple referential meaning. Through such lexical choices, the writer can align himself or herself variously with the perceived audience and the language communities in contact.

Such borrowings can be a means of expressing intentionality, defined by Myers-Scotton as "the messages conveyed by utterances in addition to those which the utterances literally denote" (1998:3). Motivation for borrowings in discourse can be analysed using parameters such as the topic and genre of the discourse, the purpose of the discourse, the identity and attitudes the writer seeks to project, and his or her relationship with the audience and with the linguistic communities in contact. In this perspective, marked borrowings reflect the situational, rhetorical or ideological dynamic of the discourse and are a manifestation of the writer in the discourse.

The presence of flagging and translations, explanations, or other metalinguistic commentary is indicative of the writer's awareness of the markedness of a linguistic variety and of its effect in a certain context. We apply the term *flagging*[2] in written discourse to include typographic markings such as italics, boldface, or quotation marks which set a word or term apart, recognizing its separate status. Typographic flagging can be purely arbitrary; in some publications, French institutional terms appear, apparently indiscriminately, in italics and roman type. But flagging also signals the writer's awareness that the borrowing is a marked choice. Quotation marks in particular are used to mark borrowings in a number of ways.[3] Palmer and Harris note that that absence of quotation marks with borrowings "indicates either a high degree of acceptance of these terms or that the writers of these documents are unconscious of these as French terms" (1990:83).

---

[2] Code-switches have been described as either *smooth* or *flagged* in studies of spoken discourse (Poplack 1993; Poplack, Wheeler and Westwood 1990).

It follows that many of the high-frequency terms in our corpus are unflagged. When we compared typographic flagging of the established Quebec English terms *vedette, poutine,* and *naive art* in *The Gazette* and in their infrequent occurrences in Canadian newspapers from outside Quebec, we found they were unflagged in *The Gazette* and flagged elsewhere, indicating an inverse relationship between flagging and frequency. Writers addressing the Quebec English community can presume that the audience is, if not fully bilingual, then at least familiar with established Quebec English borrowings (such as *dépanneur, balconville, garderie, vernissage, vedette*), which are part of a "we-code" that includes the writer and the audience and identifies with the local French community. Widely established borrowings are not necessarily consciously motivated; the more a borrowing is perceived as standard usage in the language community addressed by the discourse, the more it is unmarked and the less intentionality it conveys. "What community norms would predict is unmarked; what is not predicted is marked" (Myers-Scotton 1998:5). Myers-Scotton (1998:27) explains that speakers "know (consciously, but more often unconsciously) that the unmarked choice has occurred with more frequency than other choices in like circumstances in the community. They also have tacit knowledge that other community members share this recognition." Thus "Frequency in outcome types positively correlates with the unmarked choices" (Myers-Scotton 1998 :28).

Shared borrowings can be exploited to create complicity with the audience; borrowings in genres such as headlines, advertisements, and humour display such discursive intent. In such cases, borrowings can favour the establishment of a link of complicity and mutual recognition between the writer and audience. Checri says of borrowing in the language of advertising:

---

[3] The *Chicago Manual of Style* describes the use of italics and quotation marks to achieve special effects: these include the use of italics for "isolated words and phrases in a foreign language if they are likely to be unfamiliar to the reader" (169) as well as for "key terms in a discussion, terms with special meaning, and, in general, terms to which the reader's attention is directed" (171). As well, quotation marks are used for irony and for slang, to indicate that the word is "foreign to the normal vocabulary of the speaker" (173).

"For the meaning to be understood requires a shared experience, common cultural and ideological references....it offers the potential consumer the opportunity to grasp all the nuances of meaning, thus establishing a certain complicity between speaker and audience" (Checri 1995:62, our translation). One such example is the use of the Québécois term "Mcdo" in recent English-language advertisements for McDonald's in Quebec.

In headlines, borrowings can be used to attract attention, arouse interest, and create complicity with the audience by inviting them to share in a humorous play on words:

(4) To artists, *Oui* are the world
(from *The Gazette*, 5/11/95; headline of an article discussing the artistic community's support for the yes side in the 1995 referendum)

(5) All we are saying is give *piste* a chance
(from *The Stanstead Journal*, 10/6/98; headline of an article advocating the creation of a bicycle trail or *piste*)

(6) *Juste pour queer*
(from *The Mirror*, 27/7/1996; headline of an article about gay humour, a reference to the Montreal comedy festival Just for Laughs/Juste pour rire)

(7) It's a caisse of misfortune
(from *The Gazette*, 22/8/81, quoted in McArthur 1989:8)

Similarly, in the comic book series *Angloman* (Shainblum and Morrissette 1995, 1996), the writer creates characters such as *Poutinette, Capitaine Souche*, and *Blocman*, presuming that the parodies of the political divisions in the province will be understood by the readership.

However, the writer does not always presume that all readers will understand all borrowings. When addressing an audience perceived to be mixed or outsider or when using less frequent borrowings, the writer may act as an intermediary, flagging and explaining the borrowed term for the benefit of the unilingual reader:

(8) The new education minister, Jacques Chagnon, was considered "ministrable" – cabinet material – right from his entry into politics in 1985.
(from Don MacPherson, *The Gazette*, 15/1/94, B5)

The strategic intent may be more overt, as in the following example where the writers align themselves with the French-language community:

(9) Or was the crisis merely a heaven-sent opportunity to cynically slip us a fast one – "un gros sapin" as we say in these parts...? (from Ed Bantey, *The Gazette*, 25/1/98, A9)

(10) ...the idea of making a big fuss over it seems "quétaine", that wonderful Québécois expression that means something that is precisely between tacky and embarrassing. (from Natasha Gauthier, *The Gazette*, 23/3/97, D2)

In the promotional tourist literature in our corpus, in which the writer is presumably addressing an audience not from the local community, nonce borrowings are used to designate distinctive local features and attractions, and are often flagged and accompanied by translations, explanations or descriptions.

(11) An *aboiteau* is an earthwork dike protecting the agricultural land from saltwater infiltration. (from a tourism guide)

(12) From this point you can see how the port is protected by two piers full of "dolosses" (cement anchors which break the waves). (from a tourism guide)

Flagging is just one way in which a writer may present a term as unassimilated or unfamiliar in order to exploit its foreign status. French spelling of place names appears to be intentionally retained in tourist literature. Although Montreal and Quebec City were spelled without accents in most of the newspapers and administrative documents, *Montréal* retained its accent in 62% of its uses, and *Québec* (city), in 90% of its uses, in the touristic texts. Here the retention of the accent appears to be a rhetorical choice, designed to accentuate the exotic flavour of the destination for the would-be tourist.

Similarly, when we contrast the following noun phrases, all found in tourist brochures, we see that foreign status has been retained in (13) by retention of the head-first order of French syntax, and in (13) and (14), by retention of gender inflection.

(13) *cuisine québécoise*
(14) *Québécoise cuisine*
(15) *québécois cuisine*

Stereotypical fixed expressions such as *joie-de-vivre* and *la belle province* abound in tourist literature. This value-added use of borrowings from French has traditionally been exploited in the fields of cooking, fashion, and entertainment to endow a reference with the appeal and authenticity of a culturally different experience.

But, in the journalistic section of the corpus, a very different motivation for borrowing was observed: namely, to signal an ideological distance from the donor language community. This type of discursive intent was observed with certain high-frequency borrowings – specifically, highly charged terms of political discourse. Language and ethnic group references have heightened connotative and ideological value in Quebec's national political debate, and these associations are exploited by writers in persuasive discourse. Sometimes such borrowings were flagged: we observed a high frequency of use of quotation marks with the expressions *les autres, les anglais, maudits anglais, pure laine,* and *francization*. In such cases, quotation marks signal ironic use; like the more explicit commentary in (19) and (20), they indicate that the writer is deriding a term or distancing himself or herself from a term by associating it with the language of "the other".

(16) ...the students won't be in an entirely French environment and will therefore not be "francized"... (from *The Gazette*, 7/12/95, F1)

(17) Perhaps they also experienced what it was like to live in a society that separated the population as "pure laine" and "les autres". (from a letter to the editor, *The Gazette*, 8/5/94, B3)

(18) ...They derive their legitimacy from "le peuple québécois" and that is why "la démocratie québécoise" cannot be restrained by constitution or courts. (from William Johnson, *The Gazette*, 13/2/98, B3)

(19) ...the allophones, that dreadful word for Montreal's immigrant community... (from Alan Hustak, *The Gazette*, 30/6/96, C1)

(20) ...so-called de souche francophones... (from *The Gazette*, 30/11/94, B2)

Borrowings used in this way, whether flagged or not, are like any intralingual register shift or style shift – they are a form of variation as rhetorical strategy. The ironic use results from the fact that these borrowings represent ideological concepts with which the writer takes issue in the discourse. A borrowing cannot be isolated from its context; if it has ironic value, it is because it is associated with other indices of irony in the extended discourse. The use of borrowings with ironic or derisive intent was topic- and genre-specific in our corpus, found almost exclusively in journalistic articles and editorials on language and political issues. So we see that the loanword *francize* and its derivatives were flagged in numerous uses in journalistic texts but were not flagged at all in the administrative corpus, indicating that the word's ideological meaning is prominent in the context of the political debate but its referential meaning prevails in administrative discourse. Similarly, terms such as "pure laine" and "de souche" are used more benignly in non-political contexts:

(21) "When it comes to a pure laine classical work like Gisele, they turn shipshod." (quoted by Fee 1995:12, from a review of Les Grands Ballets Canadiens)

7 Conclusion

In this paper we have explored perspectives on motivation for lexical borrowings from French in a corpus of written Quebec English. We suggest that parameters for the interpretation of borrowing patterns include linguistic, sociolinguistic, discursive and stylistic considerations. The patterns of borrowing from French into Quebec English are indicative of the nature and sites of social, cultural, and linguistic contact between the two language communities. Borrowings from French in individual texts can reflect the ideological and rhetorical manifestation of the writer in the discourse as well as indicate relationships between the writer, the audience, and the context. Certain topics, genres and contexts are particularly conducive to intentional uses of borrowings. Intentionality appears to be linked to flagging and lack of adaptation, which indicate that the writer is treating the term as foreign to his or her repertoire, to the audience's, or to the community's.

This paper has reported on findings to date in a research project still in progress. By working with an expanded corpus and with texts from other domains and representative of a wider variety of genres, we hope in future years to gain more insight into the universal phenomenon of linguistic borrowing and the dynamics of variation in a contact situation, and specifically, in Quebec English.

References

Avis, Walter S. 1978. Some French-Canadian loanwords in Canadian English. *Walter S. Avis: Essays and articles*, 157-171. Occasional Papers of the Department of English. Kingston, Ontario: Royal Military College of Canada.

Beauchemin, Normand, Pierre Martel and Michel Théoret. 1992. *Dictionnaire des mots du français parlé au Québec*. New York: Peter Lang.

Benveniste, Emile. 1966. *Problèmes de linguistique générale*. Volumes 1 and 2. Paris: Gallimard.

Bouchard, Chantal. 1989. Une obsession nationale: l'anglicisme. *Recherches sociographiques* (Revue interdisciplinaire d'études sur le Québec et le Canada français) 30(1).67-90.

Bouchard Ryan, Ellen, Howard Giles, and Richard J. Sebastien. 1982. An integrative perspective for the study of attitudes toward language variation. In *Attitudes towards language variation: Social and applied contexts*, ed. Ellen Bouchard Ryan and Howard Giles, 1-19. London: Edward Arnold.

Bourhis, Richard. Y. 1982. Language policies and language attitudes: le monde de la francophonie. In *Attitudes towards language variation: Social and applied contexts*, ed. Ellen Bouchard Ryan and Howard Giles, 35-62. London: Edward Arnold.

Chadelet, Jean-Marc. 1995. Pour une sociolinguistique de l'emprunt lexical: l?exemple des emprunts français en anglais. *Plurilinguismes* 9-10:55-59.

Checri, Carole. 1995. L'expressivité de l'emprunt en publicité. *Plurilinguismes* 9-10:61-88.

Crystal, David. 1997. *The Cambridge encyclopedia of the English Language*. Cambridge: Cambridge University Press.

Darbelnet, Jean. 1976. *Le français en contact avec l'anglais en Amérique du Nord*. Québec: Presses de l'Université Laval.

Ede, L, and A. Lunsord. 1984. Audience addressed/Audience invoked: The role of audience in composition theory and pedagogy. *College Composition and Communication* 35(2):155-171.

Fee, Margery. 1991. Frenglish in Quebec English newspapers. *Journal of the Atlantic Provinces Linguistic Association* 15:12-23.

Fee, Margery. 1995. Using commercial CD-ROMs for dialect research: The influence of Quebec French on Quebec English in newspapers. Paper presented at 1995 Methods in Dialectology.

Fee, Margery, and Janice McAlpine. 1997. *Guide to Canadian English usage*. Toronto: Oxford University Press.

Gérard-Naëf, J. 1987. *Savoir parler, savoir dire, savoir communiquer*. Paris: Delachaux et Niestlé.

Grant-Russell, Pamela. 1998. The influence of French on Quebec English: Motivation for lexical borrowing and the integration of loanwords. Paper presented at 1998 LACUS convention; publication forthcoming in *LACUS Forum XXV*.

Grosjean, François. 1982. *Life with two languages: An introduction to bilingualism*. Cambridge: Harvard University Press.

Hamilton, Donald E. 1958. Notes on Montreal English. *Canadian Journal of Linguistics* 4:70-79.

Haugen, Einar. 1950. The analysis of linguistic borrowing. *Language* 26:210-231.

BEST COPY AVAILABLE

Keith-Ryan, Heather, and Sharon McCully. 1996. *Quebec: Bonjour, eh?: A primer for English-speakers.* Bedford, Quebec: Sheltus and Picard.

Kerbrat-Orecchioni, Catherine. 1980. *L?énonciation de la subjectivité dans le langage.* Paris: Armand Colin.

Labov, William. 1972. *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

McArthur, Tom. 1989. *The English language as used in Quebec: A survey.* Occasional Papers No. 3, Strathy Language Unit, Kingston: Queen's University.

McArthur, Tom. 1992. Quebec English. *The Oxford companion to the English language*, ed. Tom McArthur. Oxford University Press.

Maingueneau, Dominique. 1991. *L'analyse du discours.* Paris: Hachette.

Manning, Alan, and Robert Eatock. 1982. The influence of French on English in Quebec. *The Ninth LACUS Forum.* Columbia, S.C.: Hornbeam Press, 496-502.

Morsly, Dalila. 1995. El-watan, El-moudjahid, Algerie-actualités, El-djeich, Liberté, Le matin: La presse algerienne de langue française et l'emprunt à l'arabe. *Plurilinguismes* 9-10:35-53.

Mortureux, Marie-Françoise. 1997. *La lexicologie entre langue et discours.* Paris: SEDES.

Myers-Scotton, Carol. 1997. *Duelling languages: Grammatical structure in code-switching.* Oxford: Clarendon Press.

Myers-Scotton, Carol, ed.1998. *Codes and consequences: Choosing linguistic varieties.* Oxford: Oxford University Press.

Orkin, Mark M. 1971. *Speaking Canadian English: An informal account of the English language in Canada.* New York: David McKay Co.

Palmer, Joe. D. and Brigitte Harris. 1990. Prestige differential and language change. *Bulletin of the CAAL* 12(1):77-86.

Perelman, Chaïm, and Lucie Olbrechts-Tyteca. 1988. *Traité de l'argumentation.* Brussels: Éditions de l'Université de Bruxelles.

Plaice, Mary. 1984. Is that really English? *Actes du Colloque sur la traduction et la qualité de langue*, documentation du conseil de la langue française 16:67-73.

Poirier, Claude. 1985. *Dictionnaire du français québécois.* Trésor de la langue française au Québec. Laval: Presses de l'Université de Laval.

Poplack, Shana. 1993. Variation theory and language contact. In *American Dialect Research*, ed. Dennis R. Preston, 252-286. Philadelphia: John Benjamins.

Poplack, Shana, David Sankoff, and Christopher Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26:47-104.

Poplack, Shana, Susan Wheeler and Anneli Westwood. 1990. Distinguishing language contact phenomena: Evidence from Finnish-English bilingualism. In *Code-switching as a world-wide phenomenon*, ed. Rodolfo Jacobson, 185-223. New York: Peter Lang.

Reid, Gregory. 1998. Constructing English Quebec ethnicity: Colleen Curran's *Something Drastic* and Josée Legault's *L'invention d'une minorité: Les Anglos-Québécois. Post Identity*1(2):68-83.

Roberts, Roda. 1982. Frenglish, or the influence of French on English. *Les Actes du 14e colloque de l'ACLA*, 203-224.

Russell, Pamela. 1996. An investigation of lexical borrowings from French in Quebec English. *LACUS Forum XXIII*, 429-440.

Shainblum, Mark, and Gabriel Morrissette. 1995. *Angloman: Making the world safe for apostrophes!* Montreal: Nuage Editions.

Shainblum, Mark, and Gabriel Morrissette. 1996. *Angloman 2: Money, ethnics, superheroes*. Montreal: Nuage Editions.

Swales, John. 1990. *Genre Analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

University of Chicago. 1982. *Chicago Manual of Style*, 13th edition, revised and expanded. Chicago: University of Chicago Press.

Vinay, Jean-Paul, and Jean Darbelnet. 1995. *Comparative Stylistics of French and English*. Translated and edited by Juan C. Sager and M.-J. Hamel. Philadelphia: John Benjamins.

Weinreich, Uriel. 1953. *Languages in contact: Findings and problems*. The Hague: Mouton.

Département des lettres et communications
Université de Sherbrooke
Sherbrooke, Québec J1K 2R1
prussell@courrier.usherb.ca
cbeaudet@courrier.usherb.ca

Variable Article Use in Korean Learners of English

Hikyoung Lee

1 Introduction

There has been an increasing number of studies in the past few years concerning sociolinguistics and second language acquisition (Bayley & Preston 1996, McKay & Hornberger 1996 among others), the apparent reason being that these two fields complement one another to provide a holistic perspective of language use. Preston (1996:1) suggests that the relationship between the two disciplines is concerned with variation in language and the sociological and social-psychological aspects of language. In particular, variationist studies in second language acquisition have been flourishing. Variation in the "Labovian paradigm" (Ellis 1994:121, Preston 1996:2) is considered quantifiable where the frequency of occurrences are predictable according to certain terms.

The underlying assumption of this tenet is that the acquisition of language itself implies the change and variation of language skills over time. Language learning according to Labov (1996:245) requires "an understanding of probabilistic and continuously varying behavior." This varying behavior is systematic and reflects the learner's language development up to a specific stage. In this light, variation is considered a manifestation of the learners' performance and not their competence (Tarone 1988: 30).

Systematic variation in second language learners has been incorporated in the theory of interlanguage. Selinker (1969) suggests that interlanguage is an intermediate language system which is deviant from the target language but still influenced by the native language. The deviations and variations are rule governed and can be accounted for in a systematic way. Young (1991) points out that system in interlanguage is defined as a hypothetical relationship between interlanguage forms and the contexts in which they occur which may be explicitly stated and reduced to rule. He also suggests that variation may be due to surface level constraints imposed by the linguistic environment in which the forms occur.

The concept of interlanguage is greatly entwined with the Contrastive Analysis Hypothesis (Lado 1957, James 1980). This hypothesis claims that if the native language of the learner is similar to the target language, learning may be facilitated. However, if the two languages are different this may impair target language acquisition. Structures considered difficult and marked in the target language will also be continuously problematic to the non-native learner. In addition, traces of interlanguage may become fossilized and lead to persistent errors.

Errors in second/foreign language acquisition are regarded as the result of the persistence of native language interference. The Error Analysis Hypothesis (Corder 1967) analyzes errors systematically by the various criteria of origin, regularity, predictability, and variability. Errors indicate what aspects of the target language are particularly problematic to certain individual as well as groups of learners.

A common thread which can be seen in the various theoretical approaches to second language acquisition is variation. The present study will mainly be concerned with the variationist approach in examining second language acquisition.

2 The Study[1]

The present study focuses on the deletion rate of the English definite article *the* and the indefinite article *a(n)* in Korean speakers of English. The acquisition of an article system is difficult and is often imperfect. One reason may be related to the fact that articles have no clear semantic function. The article is classified as a determiner because it always signals that a noun follows. In addition, the absence of an article in many cases does not provoke severe ungrammaticality. It has been speculated that even native speakers of languages which possess an article system tend to delete articles when they do not regard them as absolutely necessary.

2.1 Learning the English Article System

In Korea, compulsory English education started when middle and high school education became compulsory in 1945. English is taught as a foreign language in Korea, which means that learners are in a non-English speaking environment. In addition, the dominant form of any foreign language teaching in Korea, is still the grammar-translation method. This implies that teaching is focused on grammatical rules as the basis for translation (Brown 1994: 52-53). In this framework, learners are trained in the passive skills of reading and writing with an emphasis on traditional grammar. However, this

---

[1] I would like to thank Gregory Guy, Kyung-Ja Park, Robin Clark, and Christine Moisset for various comments and suggestions. I also thank the audience at NWAV(E) 27 for their helpful feedback.

way of teaching minimizes the communicative aspects of the language to be learned. Although in recent years there has been a shift towards a focus on spoken language, large segments of English examinations are still devoted to translation.

In Korea, the English article system is learned in the traditional descriptive grammar framework. The use of articles is presented as a set of rules. These rules are based on prescriptive dictation. However, generalizations may seem arbitrary to the learner whose native language lacks an article system. In such languages, alternative means of expression are used in place of articles. For instance, in Korean, definite articles can be expressed by modifiers like demonstrative pronouns, and the indefinite article by numeric modifiers.

In English, the rules for article use are summarized by Fowler & Aaron (1992:748-788, 1995:727-728)[2] as follows.

(1) 1. 'The' is a definite article, it precedes a noun when the thing named is already known to the reader.
2. 'A' and 'an' are indefinite articles, they precede a noun when the thing named is not already known to the reader.
3. Use 'a(n)' with a singular count noun and do not use them with a plural noun.
4. Do not use 'the' with a plural noun or a mass noun when the noun refers generally to all representatives of what it names. Use 'the' when referring to one or more specific representatives of what the noun names.

However, these so-called 'rules' do not reflect the wide array of environments where article use is variable. In these cases, native English speakers show variable use and rely on intuition rather than generalized conventions.

2.2 The Subjects

The subjects in the present study were 49 freshmen of the Department of English at a Korean university. All subjects have had 6 years of English education from secondary schooling. The general backgrounds of the students are considered to be relatively uniform in that they all share the same native language and similar academic achievements.[3]

---

[2] *The Little, Brown Handbook* is described by the authors as being a guide to standard written English (Fowler & Aaron 1995:xvi). Although the book is mainly intended for native English speakers the most recent edition acknowledges English as a Second Language (ESL) speakers.

In a demographic breakdown of the 49 subjects, 24 were male and 25 were female. Of the subjects, 27 came from rural regions and 22 came from urban regions.

Social Factors	Rural	Urban	TOTAL
Male	17	7	24
Female	9	16	25
TOTAL	26	23	49

Table 1. Demographic information

2.3 Methodology

Data was collected from a spontaneous written composition on the topic of 'My Freshman Lab Class.' The compositions were elicited during a 50-minute time frame. The students had no previous knowledge of what the topic of the composition was, and they were not permitted to use a dictionary. A familiar topic was used in order to eliminate such factors as lack of general knowledge and insufficient English skills regarding expression. There was no restriction on the length of the composition. However, the length ranged from 1 to 2 pages. The fact that the data is in written form may suggest that the register is formal and less resilient to variation. In this sense, the written form is considered static compared to the spoken form which is dynamic. Here, written examples were elicited to test the contemplated efforts of the students and to assess errors of competence.

The MacVarb Variable Rule Program (Guy 1989) was used in this study to quantify the rate of deletion of English articles. The article errors in the compositions of the 49 subjects were analyzed and classified into those which occurred in the definite article *the* and the indefinite article *a(n)*. A total of 638 tokens were collected. This number reflects the environments where the article was deleted and where it was retained.[4] An initial run of MacVarb was conducted as well as a reanalysis of the factors in an additional modified run.

---

[3] All of the students have similar college entrance examination scores in order to have been admitted to this particular department at this particular university.

[4] Environments were determined according to Folwer & Aaron (1994, 1995)'s rules and the author's intuition.

The social variables investigated were sex and region. The variable of region indicates whether the subject attended secondary schools in a rural or an urban area of Korea. This distinction was made on the premises that region also indirectly alludes to social class to a certain degree.

(2) Social variables
    : sex
    : region (urban/rural)

The linguistic variables were formulated taking Fowler and Aaron (1992), English textbooks and frequency of appearance in the data corpus into consideration. The variables are as follows.

(3) Linguistic variables

1) Definite article THE	
: the + noun	ex) the dog
: the + adjective + noun	ex) the white dog
: the + adverb	ex) the most
2) Indefinite article A(N)	
: a(n) + concrete noun	ex) a dog
: a(n) + abstract noun	ex) an interest
: a(n) + adjective + count noun	ex) a white dog
: a(n) + adjective + mass noun	ex) a deep interest
: a(n) + adverb	ex) a little

NOTE: Based on frequency of occurrence in data corpus.

Special treatment must be given to the variable of 'article + adverb.' The classification of this variable is questionable in regard to its grammatical category. This sequence seems to appear when the following word is a superlative form of an adjective or an adverb (egs. *the most*). These instances may be considered to be idiomatic in nature and hence render the grammatical category of the superlative form as a noun. However, in order to distinguish this form from a noun, the category of 'article+adverb' is used.

2.4 Hypotheses

In this study, four hypotheses were tested. The hypotheses were posited taking all of the sociolinguistic factors into consideration.

(4) Hypotheses

Hypothesis 1

The rate of English article deletion will be considerably higher than the rate of article retention.

Hypothesis 2

Males will delete the English article more than females.

Hypothesis 3

Subjects with rural backgrounds will delete the English article more than those with urban backgrounds.

Hypothesis 4

The deletion rate of the English article will be higher if the articles precede a 'modifier+noun,' rather than directly preceding a noun.

Hypothesis 1 tests whether the deletion rate will be higher than the retention rate. This is a generalization which is formed under the rationale that the Korean language lacks an equivalent of the English article system. According to the Contrastive Analysis Hypothesis, this is considered a hindrance to acquisition. As there is no translation equivalent of the article in Korean, the subjects may be influenced by their native language and delete rather than retain this form. This form would be considered marked to them because it does not exist in their native language.

Hypothesis 2 tests differences according to sex. It is a general belief that females are more sensitive to second language forms than males. Studies in second language acquisition have proven this claim to a certain degree while in sociolinguistics studies, females are considered leaders in linguistic change (Ellis 1994:202-204).

Hypothesis 3 proposes that coming from an urban background will be more advantageous than coming from a rural one. It is widely believed that urban schools in Korea offer a higher standard of education than rural schools.[5] This holds true in relation to the differences in the socio-economic status of people living in urban areas compared to rural areas. Highschools in urban areas have a higher required score on the highschool entrance examination than rural schools. There is also a general tendency of well-qualified teachers to prefer urban schools to rural ones.

Hypothesis 4 states that an intervening grammatical category such as a modifier will influence the rate of deletion. This assumption is supported by the fact that the Korean language does not have an overt article system and incorporates this information in the use of prenominal modifiers. Therefore, if a modifier is already present in the construction the learner is assumed to be affected by native language interference and consequently not use an article. In Korean, the function of the English definite article *the* is reflected in modifiers such as *i* 'this', *ceo* 'that', and *keu* 'that (proximate to hearer)'.On the other hand, the English definite article which distinguishes singular and plural entities is expressed directly by cardinal numbers.

---

[5] Jinmin Park (p.c.1997) provided this observation. She graduated from a teacher's college and is currently a primary school teacher.

3 Sociolinguistic Variation

The results from the multivariate analyses using MacVarb will be presented in terms of the four hypotheses.

3.1 Deletion vs. Retention (Hypothesis 1)

The subjects showed a surprisingly high degree of proficiency concerning article use. This may be accounted for by the fact that the output of the subjects was in written form.

Tendency	Number of tokens	Percentage
Deletion	129	20.21
Retention	509	79.79
TOTAL	638	100

Table 2. Rates of deletion and retention of articles

The results in Table 2 show that Hypothesis 1 proved false. Of the 638 tokens collected, only 129 were instances of deletion. This indicates a mere 20.21% rate of deletion. The subjects appear to have a surprisingly high level of written language proficiency concerning article usage. Continuous exposure to grammar and written form may have sufficiently provided the setting for article acquisition. The subjects have had the chance to contemplate and correct mistakes because the samples were in written form. Therefore, they may have been consciously aware of the environments where the article must be retained.

3.2 Sex Variation (Hypothesis 2)

Turning to Hypothesis 2, the variable of sex did not prove to be statistically significant. Males showed a higher rate of deletion of the article than did their female counterparts by a difference in factor weight of only 0.12. Table 3 shows the results of the analysis.

	Deletion		
Sex	N	%	Factor weight*
Male	81	24	0.56
Female	48	16	0.44
N=129, log likelihood=-308.228			

* MacVarb results, N=number of tokens.[6]

Table 3. Sex variation in article use

3.3 Geographical Variation (Hypothesis 3)

As for Hypothesis 3, this variable was thrown out during the step up/step down analysis and thus turned out to be insignificant. This seems to be the result of the subjects having similar academic achievements regardless of geographic background in that they all entered the same university. Table 4 shows the results.

	Deletion		
Region	N	%	Factor weight
Rural	77	24	0.52
Urban	52	17	0.48

Table 4. Geographical variation in article use

As can be seen, the difference in factor weight was a mere 0.04.

---

[6] Subsequent tables will omit total number of tokens and log likelihood.

3.4 Grammatical Variation (Hypothesis 4)

The most important finding of the present study was the significance of the grammatical environment immediately following the target article. A low deletion rate occurred when the article directly preceded the noun it was modifying. Table 5 shows the results.

		Deletion		
Article	Following categories	N	%	Factor weight
the	+ noun	39	14	0.34
	+ adjective + noun	44	24	0.50
	+ adverb	6	33	0.60
a(n)	+ concrete noun	8	18	0.41
	+ abstract noun	12	31	0.58
	+ adjective + count noun	10	26	0.52
	+ adjective + mass noun	7	44	0.70
	+ adverb	3	14	0.34

Table 5. Deletion rates according to following categories

The variables showed mixed results. A slight tendency was apparent in that the subjects seemed to delete the article more when it preceded an adjective or an adverb than when it preceded a noun. This provides evidence for Hypothesis 4. The subjects also deleted more when the following noun was an abstract noun than when it was a concrete noun. This triggering of deletion may be due to the notion of the general use of the indefinite article as an indicator of singular concrete nouns. In this sense, it may be difficult for the second language learner of English to relate the indefinite article to abstract entities.

3.5 Reanalyses

Next, the initial run was modified so that some of the linguistic constraints were conflated. Adjustments were made to see whether a significant difference existed according to whether the article was definite or indefinite. In addition, article use in general and the effects of linguistic environments were examined. The new adjustments in the factor groups showed varied results.

3.5.1 Type of Article

The first reanalysis micro-analyzed the article according to type. The results showed that there was no significant difference in distinguishing between the definite article *the* and the indefinite article *a(n)*.

	Deletion		
Type	N	%	Factor weight
Definite 'the'	89	19	0.46
Indefinite 'a(n)'	40	25	0.54

Table 6. Deletion rates according to type of article

The subjects appeared to be prone to deletion of the article regardless of its particular grammatical usage. This finding also somewhat indirectly indicates that the subjects view articles in a general manner and may unconsciously group them as one category.

3.5.2 Modified Grammatical Environments

The second reanalysis conflated the grammatical environments following the article as well as the type of article. In particular, the distinction between count nouns and mass nouns was disregarded.

	Deletion		
Grammatical order	N	%	Factor weight
article + noun	59	16	0.40
article + adjective + noun	61	26	0.55
article + adverb	9	23	0.54

Table 7. Deletion rates of general article use

The results indicate that in general article use the lowest deletion rate is seen when the article directly precedes a noun. The finding reinforces Hypothesis 4, as the deletion rates were found to be higher when the article precedes an adjective or an adverb.

4 Summary and Conclusions

The data examined in the present study showed the various sociolinguistic factors which affect the deletion of the definite and indefinite article in Korean learners of English. The overall low article deletion rate seems to suggest that the subjects have acquired the article system to a large degree. Errors are made when the use of the article is arbitrary and obscure.

The results also showed that only one of the four research hypotheses posited were confirmed.

(5) Hypotheses testing results

Hypothesis	Proposal	Status
Hypothesis 1	deletion > retention	disconfirmed
Hypothesis 2	males ≈ females	disconfirmed
Hypothesis 3	rural > urban	disconfirmed
Hypothesis 4	+ modifier + noun > + noun	confirmed

Deletion of both the definite and indefinite article appears to be systematic and sensitive to both internal and external factors. Males showed similar rates of deletion as females. There was no apparent significance regarding the geographical background of the subjects. In addition, there was no substantial difference in the deletion rates of the definite and indefinite article.

Significant differences were found in the linguistic environment the article appeared in. There was a tendency of higher deletion when the article preceded an adjective or an adverb than when it directly preceded a noun. This can be explained by native language transfer, as the role of the article is played by modifiers in the subjects' first language of Korean. Since the position where an article should appear is already occupied with a modifier, the subjects might have overlooked the fact that an article might be necessary in this context. In the case of the indefinite article, the deletion rate was higher when the article preceded an abstract noun than when it preceded a concrete noun. This may be explained by the nature of concrete nouns as being relatively more countable than abstract nouns.

It is concluded from the present study that the phenomena of article deletion in Korean learners of English seems to be greatly influenced by the learners' native language. The study has thus attempted to incorporate variationist methods in order to unveil the nature of article use in the written language of Korean learners of English.

References

Ahn, Byung-Kyoo. 1992. 'Native speaker perception versus writer intent in errors in Korean students' written English discourse.' *Language Research*, 28(1). Seoul: Language Research Institute, Seoul National University. p.174-193.

Bayley, Robert. & Preston, Dennis.R. 1996. *Second Language Acquisition and Linguistic Variation*. Philadelphia: John Benjamins.

Brown, H.Douglas. 1994. Teaching by Principles: An Interactive Approach to Language Pedagogy. Englewood Cliffs, NJ: Prentice Hall Regents.

Corder, S.P. 1981. *Error Analysis and Interlanguage*. London: Oxford University Press.

Ellis, Rod. 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Fowler, H.Ramsey. & Aaron, Jane E. 1992. *The Little, Brown Handbook*. (5th Ed). New York: HarperCollins College Pub.

Fowler, H.Ramsey. & Aaron Jane E. 1995. *The Little, Brown Handbook*. (6th Ed.) New York: HarperCollins College Pub.

James, Carl. 1980. *Contrastive Analysis*. Essex: Longman.

Lado, Robert. 1957. *Linguistics Across Cultures*. Ann Arbor: The University of Michigan Press.

Lee, Hikyoung. 1993. *A Study on Korean EFL Learners' Composition Errors*. Unpublished master's thesis, Korea University, Seoul, Korea.

McKay, Sandra L. & Hornberger, Nancy H. 1996. *Sociolinguistics and Language Teaching*. New York: Cambridge University Press.

Park, Kyung-Ja. 1992. 'A contrastive analysis of Korean EFL learners at the university level.' *Phoenix*, 31. Seoul: Korea University Press. p.1-65.

Preston, Dennis R. 1989. *Sociolinguistics and Second Language Acquisition*. Great Britain: Oxford University Press.

Preston, Dennis R. 1993. 'Variation linguistics and SLA.' *Second Language Research* 9(2). p.153-172.

Richards, Jack C. 1974. *Error Analysis*. London: Longman.

Santos, Terry. 1987. 'Markedness theory and error evaluation: An experimental study.' *Applied Linguistics*, 8(3). p.207-218.

Selinker, Larry. 1992. *Rediscovering Interlanguage*. New York: Longman.

Tarone, Elaine. 1988. *Variation in Interlanguage*. London: Edward Arnold.

Young, Richard. 1991. *Variation in Interlanguage Morphology*. New York: Peter Lang Publishing, Inc.

Young, Richard. 1996. 'Form-function relations in articles in English interlanguage.' In R. Bayley and D. Preston, *Second Language Acquisition and Linguistic Variation*. Philadelphia: John Benjamins. p.135-176.

Department of Linguistics
University of Pennsylvania
619 Williams Hall
Philadelphia, PA 19104-6305
hikyoung@babel.ling.upenn.edu

The *University of Pennsylvania Working Papers in Linguistics (PWPL)* is an occasional series produced by the Penn Linguistics Club, the graduate student organization of the Linguistics Department of the University of Pennsylvania.

Publication in this volume does not preclude submission of papers elsewhere; copyright is retained by the authors of the individual papers.

Volumes of the Working Papers are available for $15, prepaid. Please see our web site for additional information.

The PWPL Series Editors

Jim Alexander
Alexis Dimitriadis
Na-Rae Han
Elsi Kaiser
Michelle Minnick Fox
Christine Moisset
Alexander Williams

How to reach the PWPL

U. Penn Working Papers in Linguistics
Department of Linguistics
619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104–6305

http://www.ling.upenn.edu/papers/pwpl.html
working-papers@ling.upenn.edu

The Loss of Auxiliary Selection in English

Mimi Lipson

1 Introduction

At some point in its history, English ceased to be a language that, like most Germanic and Romance languages, uses the HAVE auxiliary with some verbs and the BE auxiliary with others to form the Perfect. This project is a corpus-based investigation of the decline of auxiliary selection in English. A large corpus of sentences was assembled from literary texts available on the Internet through Project Gutenberg[1], covering the years 1560–1875 (= year of author's birth). The results show that, during the long period of variable auxiliary selection, the choice of auxiliary reflected a previously un-encoded contrast in meaning.

Evidence for semantic structure is always indirect, being more a matter of what entailments go with a particular construction than what constructions are possible or grammatical in a language. This fact makes the diachronic study of semantics through corpora particularly difficult. In this case, both assembling a usable corpus and interpreting the pattern of variation required some basis for characterizing the relevant semantic qualities of the sentence-level tokens. As far as I know, projects such as these are uncommon; therefore, the second and third sections focus on the methodological issues that came up when assembling and analyzing the corpus.

A major theme of contemporary sociolinguistics is the linguistic interpretation of variation, and specifically, how the distribution of new and old forms can be used to answer questions about how innovative forms embed themselves in the grammar. Toward that end, I propose a linguistic trajectory of change for the present case in Section 4. The implications of this embedding mechanism will be discussed in Section 5, particularly with respect to the *competing grammars* model of linguistic change developed in the work of Kroch and others. Finally, there is a growing body of synchronic work on the variable distribution of HAVE and BE in the Germanic and Romance languages (Platzak 1987, Kayne 1993, Freeze 1992, Iatridou 1995), and in the last section I will say a few words about what I believe to be the relevance of this work to the diachronic situation in English. (Readers interested in a more formal semantic treatment of this particular case are referred to Lipson 1999).

---

[1] e-texts: http//www.promo.net/pg/list.html

2 Auxiliary Selection

In auxiliary selection languages, the potential BE-selecting verbs are unaccusatives—that is, verbs in which the single argument shows object-like properties, as in (1), as opposed to subject-like properties, as in (2).

(1) a. Han *har/er rejst
*he *has/is gone* (Danish)
b. Si *ha/è arrivati
*they *have/are arrived* (Italian)
(2) a. Han har/*er sovet.
*he has/*is slept*
b. Si ha/*è mangiati
*they have/*are eaten*

There are many internal lexical-semantic distinctions within the general class of unaccusative verbs (for a thorough discussion of unaccusative verb classes, see Levin and Rappaport-Hovav 1996), and auxiliary-selecting languages differ in how widespread BE-selection is among the unaccusative verb types.

3 The Gutenburg Corpus

3.1 Sampling Procedures

In order to conduct a constrained and systematic search of texts from various historical periods, I restricted my search to the class Levin and Rappaport-Hovav call 'verbs of inherently directed motion,' a class which includes verbs such as *come, go, fall, rise/arise, depart, arrive, land, return, gather,* and *meet.* It was necessary to maintain some degree of comparability between verbs, since we have no other assurance that the individual verbs within the class are not behaving differently (through some form of classificational drift or lexical diffusion). Of the verbs sampled, only *come, go,* and *fall,* and *become*[2] occur in sufficient numbers for cross-verb comparison, and so only these verbs were included in the corpus.

The object in assembling the corpus was to collect all tokens of Perfects within this verb class in order to study the decline over time of the use of BE + past participle as a Perfect. However, contemporary usage suggests that not all such strings are actually Perfects. In particular, consider the case with *gone*, which occurs in construction with both auxiliaries:

---

[2] Though *become* is transitive, it is universally classed as a BE verb in auxiliary-selecting languages.

(4) a. John has gone
    b. John is gone

The alternation in (4) does not mean that there is (variable) contemporary BE selection when forming a Perfect of the verb *go*. There are three types of evidence against treating (4b) as an alternative Perfect construction. First and most obviously, there are many types of modification which, in contemporary English, are restricted in their use with *gone*, several of which are listed in (6); specifically, the construction does not combine comfortably with modification that delimits the path, goal, beginning time, or manner of going.

(6) a. John had gone to the store
    b. ?John was gone to the store *(goal)*
    c. John had gone on the interstate
    d. ?*John was gone on the interstate *(path)*
    e. John had gone as soon as possible
    f. ?John was gone at 12 sharp *(beginning time)*
    g. John had gone effortlessly/by car
    h. ?John was gone effortlessly/by car *(manner)*

Also, BE + *gone* and HAVE + *gone* have different entailments with respect to the current status of the underlying eventuality: (4b) entails that John is still gone, whereas (4a) carries no such entailment: John may have gone and then returned. In this sense, *gone* resembles past participles like *drunk* or *finished*:

(7) a. John had finished (with) his homework
    b. John was finished with his homework
(8) a. John had drunk
    b. John was drunk

This distinction turns out to be of little help in the sampling process, since it is not usually possible to deduce these sorts of temporal entailments from a written text. However, I discuss below a way in which the sample seems to reflect the entailment facts of contemporary HAVE/BE alternations.

The most serious objection to analyzing (4b) as a BE-selecting Perfect is the fact that a Perfect may be formed from BE + *gone*:

(9) John *has been gone* for ages.

There is evidence that this sort of stative adjective used to be more widespread within the verb class in question, for instance:

(10) And with that Diabolus gave back, thinking that more aid *had been come.* (Bunyan)

For this reason, it would not have been sufficient to simply remove the *gone* tokens from the sample. I needed some way of ensuring that these stative adjective uses were not counted in the sample—i.e., that all the tokens included were unambiguous Perfects. Essentially, I used as a guide my own intuitions about what can currently modify BE + *gone* in contemporary speech. I included in my sample tokens occurring with any hint of agentivity—mainly purpose, as in (11), delimiting PPs such as those in (12), and certain kinds of adverbials, as in (13), as well as tokens coordinated with other verbal past participles, as in (14):

(11) a. They are come to give us joy.
  b. I am come on purpose to quarrel with you
  c. She was come in herself for the Stilton cheese,
(12) a. They were gone to Hartfield. *(goal)*
  b. I was really gone from Randalls. *(source)*
  c. Is Mr. Elton gone on foot? *(instrument)*
  d. I had not gone three steps... *(extent)*
  e. No sooner had I come to the point... *(result)*
  f. ...after so many hazards as I had gone through... *(path)*
(13) She had come as fast as she could *(rate/pace)*
(14) When they had been all walking together, he had so often come and walked by her, and talked so very delightfully!—

All tokens that did not have some type of disambiguating modification were excluded from the sample.

3.2 Sample Overview

In order to get an idea of when the HAVE Perfects first began to appear, I looked at the Penn-Helsinki corpus of Middle English, which is comprised of texts dated from AD 1150 to AD 1500. The entire corpus contained only fourteen examples of HAVE + *come.* Of these, twelve occurred in irrealis (i.e., non-factive) environments—either in counterfactuals, as in (15), or as complements of modal verbs, as in (16):

(15)...and *had I not that tyme comen* he sholde haue taken his lyf from hym...
(16)...for south ye *myght have comen* to my counter, ...

The Middle English corpus established two things: 1. that an adequate sample would have to cover a later period of the language, and 2. that irrealis uses were the first places where BE was replaced with HAVE in this verb class. Indeed, looking at the Early Modern English data, it appears that have was categorically used with modals and counterfactuals, except where be was used to denote futurity. Since they did not show variability, the modal and counterfactual uses were also eliminated from the sample.

Table 1 shows the token distribution of the four verbs in the whole sample as a whole, divided by the period during which the author was born. Table 2 shows the percentages of be Perfects for the four verbs, and for a composite class of the three transitive verbs (*come, go,* and *fall*).

When plotted graphically, the decline in the percent of BE Perfects in the sample follows an S-shaped curve[3], as in Figure 1.

All Perfects	1560-1575	1608-1625	1660-1675	1710-1725	1760-1775	1810-1825	1860-1875	Totals
come	39	96	30	19	54	144	23	405
gone	76	38	72	17	59	83	33	378
fallen	21	11	17	7	25	32	16	129
become	28	9	10	9	27	37	14	134
Total	164	154	129	52	165	296	86	1046

Table 1: Project Gutenberg texts (by author's year of birth)

BE Perfects	1560-1575	1608-1625	1660-1675	1710-1725	1760-1775	1810-1825	1860-1875
come	32	93	20	5	19	26	0
gone	49	25	40	6	19	15	2
fallen	15	8	7	2	0	2	0
become	28	9	10	7	11	2	0

Table 2: BE Perfects in the Gutenberg corpus

---

[3] The period of initial increase is probably due to stylistic differences between the earliest texts, which drew heavily from the morality plays of Bunyan, and the second-earliest category, in which the plays of Shakespeare are most heavily represented.

BEST COPY AVAILAB

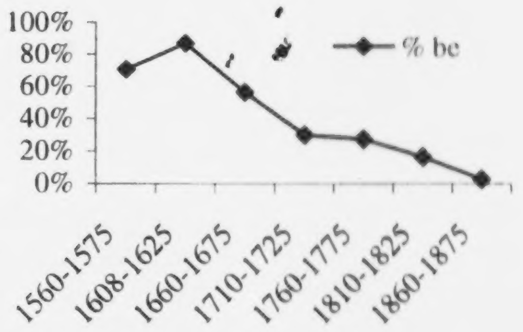


Figure 1: Rate of be-perfects with *come, go,* and *fall.*

This trajectory is a typical quantitative picture of the loss of an outgoing variant: in case after case of syntactic, morphological, or phonological change, the rate of change is slow at the beginning and end of the process, and slower in the middle. It is interesting to note that the pattern holds in a linguistic change which, as we will see, involves meaning differentiation.

When the verbs are plotted individually, as in Figure 2, it is apparent that they pattern as a class—with the exception of *become* which, being a transitive verb, is arguably not a real member of the class.

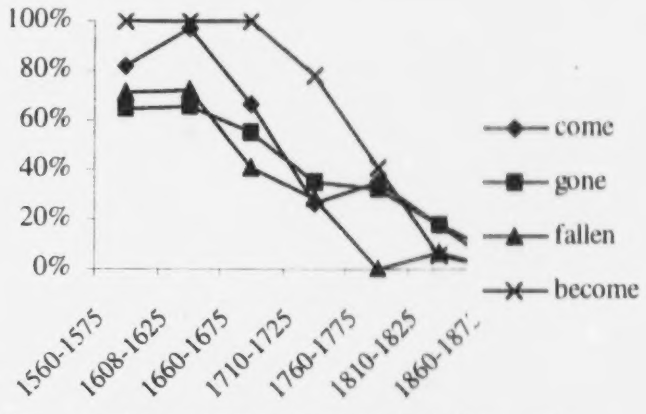


Figure 2: Rates of BE perfects for individual verbs

4 Semantic Differentiation of the Auxiliaries

The earliest data show that modal/counterfactual use of HAVE Perfects were the first to appear in this verb class. In other words, these irrealis contexts seem to have been the wedge that introduced the HAVE Perfect to this verb class. The question then becomes: how did the HAVE Perfect spread to non-irrealis contexts? As I worked with the Gutenberg texts, I kept noticing examples such as those in (17).

(17) a. For *after* that the children of God had gone in vnto the doughters of men and had begotten them childern, the same childern were the mightiest of the world and men of renowne (Bunyan).
b. Thus *twice before*, and jump at this dead hour, with martial stalk hath he gone by our watch (Shakespeare).
c. For a mile or thereabouts my raft went very well, only that I found it drive a little distant from the place where I had landed *before* (Defoe).
d. ...and my conscience, which was not *yet* come to the pitch of hardness to which it has since, reproached me with the contempt of advice, and the breach of my duty to God and my father (Defoe).
e. Am I come near ye *now*? (Shakespeare)

I began to strongly suspect that, during the period when there was variation in the auxiliary used to form the Perfect in this verb class, the morphological variation was exploited semantically. Roughly, the two forms seemed to show a tendency toward the entailment difference pointed out above in the discussion of contemporary *have gone/be gone:* BE kept appearing in contexts where the eventuality persists, as in (17d and e), where the sentence is modified by *yet* and *now*. Conversely, HAVE seemed to appear with unusual frequency where the eventuality is more certain to have ended or taken place entirely in the past, as in (17a, b, c), where it coincides with the modifiers *after, twice,* and *before*.

4.1 Semantic Heterogeneity of the Perfect

The meaning distinction alluded to is roughly in line with the distinction between what are sometimes referred to as the Universal Perfect and the Existential Perfect (cf. Dowty 1979). In pseudo-Reichenbachian terms, where T=topic time (time during which the event or state referred to took place or held), R=reference time (e.g. now for present Perfects; some time in the past for past Perfects), and S=speech time (the time of utterance), the distinction is whether or not T and R are ever contemporaneous: the Universal Perfect means that the reference time is included in the interval during which the eventuality holds. The distinction is illustrated in (18,19), with the box notation indicating the duration of the topic time.

(17) Mary had always lived there.

Universal: R... S
Mary lives there

(18) Mary had read Valley of the Dolls.

Existential: Mary reads R...S

I use the terms Universal and Existential with caution, since the Universal Perfect is said to be available only with stative verbs (but see below for a sense in which the BE Perfect behaves like a Perfect formed with a stative verb); however, it does seem that during the period of auxiliary variation, the auxiliaries were (or at least tended to be) used to encode the distinction in terms of the topic time/ reference time relation. This is a difficult point to argue when dealing with historical texts, since the evidence for these sorts of temporal entailments is quite indirect. However, as suggested above, there are some differences in the way these two Perfects can be modified; for instance, the Universal Perfect cannot be modified by iterative adverbials, like *many times* or *twice*. Table 3 shows that HAVE is more prevalent than would be expected (given the overall distribution) with modification that contraindicates the Universal Perfect reading, whereas BE is more likely with *now*.

Year of birth:	1560-1575	1608-1625	1660-1675	1710-1725	1760-1775	1810-1825
Now + BE	20	22	13	7	17	4
Now + HAVE	0	1	0	2	2	10
Times* + BE	5	4	0	0	0	0
Times + HAVE	8	8	1	6	3	4

Table 3: Effects of temporal modification
* = X *times, before, since, after*

When plotted as percentages of Perfects selecting BE (Figure 3) and HAVE (Figure 4), the effect is quite clear: though the direction and general rate of change is consistent in every category, the overall rates are dramatically different.

The HAVE/BE distribution is reminiscent of the *be gone/have gone* alternation in another way: because only *be gone* is durative in nature, *have gone* cannot occur in a *while* clause:

(20) Mary read a book while John was/*had gone

The problem is not that Perfects cannot be temporal antecedents for *while*, but rather that the underlying eventuality must be stative, as in *know* and *have:*

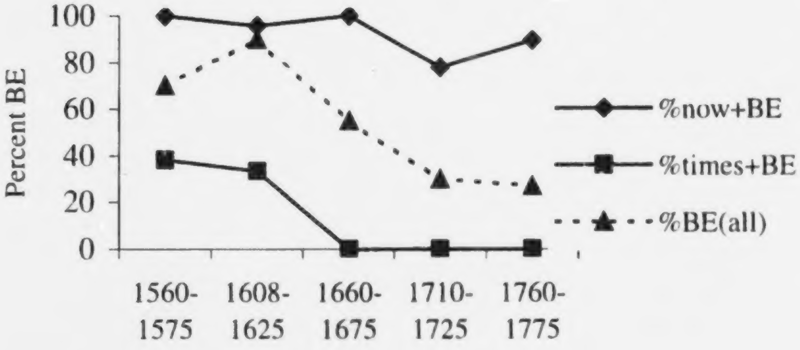


Figure 3: % BE with temporal modifiers

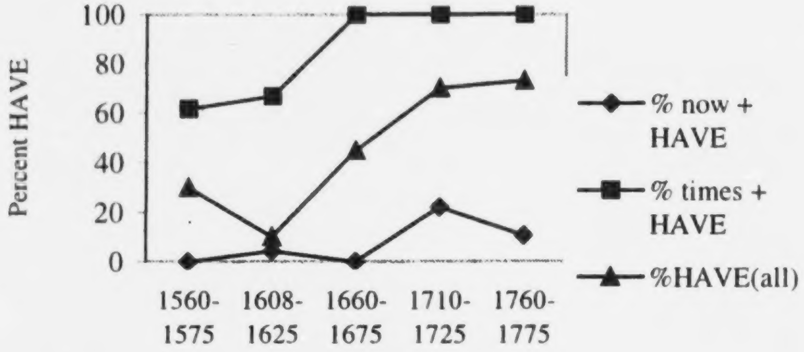


Figure 4: % HAVE with temporal modifiers

(21) a. I've read *Valley of the Dolls* four times while I have lived here.
b. John has read *Valley of the Dolls* four times while I have known him.

There are several examples of BE + past participle *while* clauses in the corpus, and some with modification that identifies the construction as a Perfect, for instance:

(22) a. We found her a ship of Bristol, bound home from Barbadoes, but had been I own out of the road at Barbadoes a few days before she was ready to sail, by a terrible hurricane, *while the captain and chief mate were both gone on shore*. (Defoe)
b. The dwarf, watching his opportunity, *while Glumdalclitch was gone to the side-board*, mounted the stool that she stood onto take care of me at meals, took me up in both hands, and squeezing my legs to-

> gether, wedged them into the marrow bone above my waist, where I stuck for some time, and made a very ridiculous figure. (Swift)

Examples like these suggest that the class of verbs has been reclassified in terms of the possible aspectual interpretations they support.

4.2 A Semantic Trajectory

The question now becomes: how did the HAVE Perfect come to be used in non-irrealis contexts? It seems to me, intuitively, that it is not a coincidence that the HAVE Perfects, which started out in irrealis environments, next occurred preferentially in Existential Perfects. So, what do counterfactuals and modals have in common with the Existential Perfect sketched in (19) but not the Universal Perfect sketched in (18)? In both cases, the eventuality is *not* asserted to persist into the present. We will see if we can make some use of this fact.

Iatridou (1996) accounts for the contribution of past-tense morphology in the interpretational contrast between normal conditionals (23a) and counterfactuals (23b):

(23) a. If you live here, you are home now.
  b. If you lived here, you would be home now.

The difference, essentially, is that (23b) is anti-factive: it cannot be felicitously uttered if in fact you *do* live here—whereas (23a) is factively neutral. Iatridou proposes that the link between the use of past morphology for *non-present* denotation and for *counterfactual* denotation is the relational quality of EXCLUSION (or, more accurately, non-inclusion). This relation can be asserted of time intervals (as in the simple past denotation) or of sets of worlds (the modal interpretation). In other words, the use of past tense in a conditional like (23b) triggers the inference that no world in which the proposition "you live here" is true is included in the set of worlds that are consistent with what we know to be consistent with the present one.

There is a similar sort of interpretational side-effect connecting irrealis contexts to the punctual, Existential interpretation of the Perfect (and to the interpretation of *have gone* as opposed to *be gone*): In the Existential interpretation, the set of time intervals in which the eventuality holds (i.e., the Topic Time) is not included in the time of reference; this is not the case with the Universal Perfect. In the case of the irrealis context, this relation holds between the possible worlds evoked by the modal/counterfactual and the actual world at the time of speech: *John might/could/should have gone* does not entail that John did in fact go.

The connection sketched above assumes that a relation can be extended from one semantic domain to another. In the case of counterfactuality and past marking, a morphological-semantic mapping applied in the unmarked case in temporal structures is superimposed upon a modal structure. The trajectory I am suggesting is an extension of the environments selecting the HAVE auxiliary from the modal domain to the aspectual domain. Note that this would not be unprecedented: there are some other examples of languages where the Perfect morphology is used to denote special epistemic positions, such as the Macedonian Dubitative Perfect (Friedman 1988) and the Bulgarian Perfect of Evidentiality (Izvorski 1997).

5 Competing Grammars?

There is a steady accumulation of research on the synchronic facts of auxiliary selection. Attempts to explain the distribution of auxiliaries, either within a language or crosslinguistically, have tended toward syntactic explanation—such as the accounts of Kayne (1993) and Freeze (1993), both of whom propose that the HAVE auxiliary is the spellout of a movement-driven incorporation of BE with an abstract pronoun-like feature, or Platzack (1987) who argues that auxiliary selection among the Scandinavian languages is an epiphenomenon of the null subject parameter. Whatever the merits of the syntactic approach to auxiliary selection, it is difficult to see how such analyses could explain the diachronic situation in English: the decline in auxiliary selection did not coincide with any change in pro-drop or word-order possibilities. The next obvious place to look for an explanation of the loss of this (putatively) syntactic feature is the work of historical syntacticians.

In his writing on historical syntax, Kroch (1994) invokes the Blocking Principle from morphology as a cause of change: *Stable linguistic systems do not permit equivalent morphological (or syntactic) doublets* (i.e., dived/dove). He argues that all morphosyntactic change indicates the existence of competing grammars, with each member of the doublet represented in one or the other grammar. The argument for competing grammars has been convincingly made in cases such as the evolution of periphrastic *do* in Middle English (Kroch 1989) or the change from V-to-C to V-to-I movement in Yiddish (Santorini 1993), and on the evidence of cases such as these, Kroch (1994) makes some explicit claims about the underlying mechanism of linguistic change: *"We have seen that the historic evolution of competing variants in syntactic change is similar to the evolution of morphological doublets. In both cases, the coexistence of variant forms is diacronically unstable: one form tends to drive the other out of use and so out of the language" (Kroch 1994: 17).* The old grammar is replaced by a new one at the

BEST COPY AVAILAB

same rate in all environments, since each environment is representative of an underlying grammar which is the actual locus of change. In some cases, it can appear that certain environments "lead" in the change, since the input rate of application of the new form may be greater or less, but this is not to be taken as evidence that change progresses independently in individual linguistic environments. In other words, Kroch argues that it is *not* the case that forms in specific environments change by a process of analogy with other environments—i.e., "generalization" of the application of a new rule.

Since auxiliary selection has so often been analyzed as a syntactic phenomenon, we would expect that this sort of analysis, and specifically the invocation of the Blocking Principle, would apply to a case where auxiliary selection is lost in a language. However, note that the Blocking Principle, as stated, can be satisfied in more than one way: either one or the other morpheme may be lost *or* their status as doublets may change via a process of semantic differentiation. In the present case, it seems clear that the process involves more than a meaning-neutral transition from one morpheme (BE) to another (HAVE). Along the way, there seems to have been a language-internal redistribution of grammatical distinctions. Of course this begs the question: why *did* the BE Perfect go out of use? I don't have an answer for this, but possibly it was a problem of dual interpretations of past participles like *fallen* and *become*. In other words, perhaps having the resources to distinguish Existential Perfects from Universal Perfects in only one verb class was too cumbersome and asymmetrical. Note again that meaning one entails the other: *I am come* entails *I have come*. Therefore, any situation in which *I am come* is felicitously uttered also supports the statement *I have come*. Therefore, if one or the other interpretation had to survive, it makes sense that it would be the more broadly applicable.

6 Conclusion

There is an accumulating body of examples where morphosemantic change involves language-internal morphological shifts from one semantic domain to another. The existence of such a path of change presents a challenge to the universal applicability of a competing grammars analysis of language change and suggests that a general model of morphosemantic change will require a somewhat different outlook—one more sensitive to the possibilities of endogenous change.

This particular case study leaves many unanswered questions. The details of auxiliary selection may be of limited interest to sociolinguists, but I believe the more general questions are worth asking: how can morphosemantic change be most efficiently investigated in a corpus? Is morphosemantic change like other sorts of linguistic change? What can diachronic variation tell us about synchronic variation, and vice-versa?

References

Anagnostapoulou, Elena, Sabine Iatridou and Roumyana Izvorski. 1997. On the Morpho-Syntax of the Perfect and how it relates to its Meaning. Unpublished MIT/Penn manuscript.

Dowty, David. 1979. *Word Meaning in Montague Grammar*. Dordecht: Reidel.

Freeze, Richard. 1992. Existentials and Other Locatives. *Language*

Friedman, Victor A. 1988. The Category of Evidentiality in the Balkans and the Caucasus American Contributions to the Tenth International Congress of Slavists. Schenker, Alexander M. (ed.). 121-139.

Iatridou, Sabine. 1995. To Have and Have not: On the Deconstructionist Approach. WCCFL 14.

Iatridou, Sabine. 1996. The Grammatical Ingredients of Counterfactuality. ms. University of Pennsylvania.

Izvorski, Roumyana. 1997. The Present Perfect as an Epistemic Modal. Paper presented at SALT 7.

Kayne, Richard. 1993. Towards a Modular Theory of Auxiliary Selection. *Studia Linguistica.*

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1: 199-244

Kroch, Anthony. 1994. Morphosyntactic Variation. In K. Beals et al., eds. *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasessions on Variation and Linguistic Theory.*

Labov, William. 1994. *Principles of Linguistic Change. Vol. 1: Internal factors.* Oxford & Cambridge MA: Blackwell.

Levin, Beth and Malka Rappaport-Hovav. 1996. *Unaccusativity.* Cambridge, MA: MIT Press

Lipson, Mimi. 1999. *Co-variation of form and meaning in the loss of auxiliary selection in English.* Paper presented at DGFS, Workshop *1 (Meaning Variation-Meaning Change)*

Platzack, Christer. 1987. Have/Be Variation in the Scandanavian Languages.

Santorini, Beatrice. 1993. The Rate of Phrase-structure Change in the History of Yiddish. *Language Variation and Change, 5:257-283.*

Weinreich, Uriel, William. Labov and Marvin I. Herzog. 1968. Empirical Foundations for a Theory of Language Change. In W. P. Lehmann and Yakov Malkiel, eds. *Directions for Historical Linguistics, 95-188.* Austin & London: University of Texas Press.

Department of Linguistics
University of Pennsylvania
Philadelphia, PA 19104-6305
lipson3@ling.upenn.edu

The *University of Pennsylvania Working Papers in Linguistics (PWPL)* is an occasional series produced by the Penn Linguistics Club, the graduate student organization of the Linguistics Department of the University of Pennsylvania.

Publication in this volume does not preclude submission of papers elsewhere; copyright is retained by the authors of the individual papers.

Volumes of the Working Papers are available for $15, prepaid. Please see our web site for additional information.

The PWPL Series Editors

Jim Alexander
Alexis Dimitriadis
Na-Rae Han
Elsi Kaiser
Michelle Minnick Fox
Christine Moisset
Alexander Williams

How to reach the PWPL

U. Penn Working Papers in Linguistics
Department of Linguistics
619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104–6305

http://www.ling.upenn.edu/papers/pwpl.html
working-papers@ling.upenn.edu

Syntactic Change in Progress: Semi-Auxiliary *busy* in South African English[1]

Rajend Mesthrie

"I'm busy going crazy."
Harassed graduate student of Linguistics, University of Cape Town.

1 Introduction

A prominent lexico-grammatical element in South African English (henceforth SAE) is *busy*, used together with the usual *be + -ing* of standard English to mark ongoing activity. Its existence on the borderline of lexis and grammar can be seen from its being just about the only grammatical marker to have an entry in the *Dictionary of South African English on Historical Principles* (DOSAEHP). The entry in the earlier fourth edition of the *Dictionary of South African English* (DOSAE) is as follows (edited slightly for illustrative purposes):

> **busy** adj. substandard. Used redundantly in SAE with an *-ing* form of certain forms to indicate progressive, signifying "engaged in" or equivalent. [translation *besig (om te)* which in Afrikaans verb phrases expresses the 'continuous present' or progressive aspect of a verb].
>
> I rushed in and found the two infants busy having convulsions - as though there were not enough troubles. K. McMagh *Dinner of Herbs* 1968.
>
> There was a sign on the kitchen door saying 'Do not disturb me, am busy praying. *Sunday Times* 28.10.90.[2]

This phenomenon was first examined in some detail a decade ago by Lass and Wright. Their characterisation (1986:213-4) was as follows:

---

[1] Earlier versions of this paper were read at the Linguistics Society of Southern Africa Conference (Potchefstroom, July 1998) and N-WAVE (October 1998). My thanks go to the participants for helpful comments and to the following persons for passing on their data to me: P. Silva and L. Pienaar (Dictionary Unit, Rhodes University) & L. Pierce and C. Jeffery (The Corpus of South African English).

[2] DOSAE has *disturbed* not *disturb*. I have resisted the [*sic*] for this typo, which would distract from the point being made.

> In [SAE] (as in Afrikaans) for instance, one may be described as being busy waiting, thinking, wondering, quite appropriately. This suggests that *busy* functions simply as a progressive marker, without the restriction that the lexical verb denoting the activity be overtly busy. The overall sense of the construction in [SAE] is minimally different from what we might term the simple progressive as in *I am relaxing* and *I am working*. The busy construction, if anything, emphasizes that the subject is in the process of, or in the midst of, or occupied with the experience of relaxing or working. It then specifically excludes any suggestion of endpoints.

I have little or no modification to this characterisation, but would like to attempt a more detailed appraisal of the contexts which favour *busy* in the light of further data. The data base of this study is a varied one, comprising the following:

(a) Material gleaned from current newspapers and (to a lesser extent) radio and television.
(b) Sentences from the speech of Capetonians in a variety of settings which were diarised with as much social and contextual background information as possible.
(c) Sentences from the published dictionaries of the Dictionary Unit of Rhodes University (DOSAE and DOSAEHP), as well as unpublished sentences from their corpus, kindly made available to me.
(d) Letters housed in the Cape Archives, written by first generation British settlers in the eastern Cape to the governor in Cape Town.
(e) My own detailed corpora pertaining to contemporary spoken Black and Indian English of South Africa.

In addition, I have begun compiling an international comparative corpus on *busy*, based on the Corpus of South African English (originally part of the International Corpus of English - ICE) and on available British and U.S. corpora. Since this work is still in progress, it will not be reported here.

2 A Historical Perspective

Lass and Wright were concerned with making a general point about the relative influences of endogenous developments in language change versus contact phenomena. In particular, they showed that while *busy* was usually assumed to be calqued on the Afrikaans progressive *besig om te* + V ('*busy* *to* + V'), the history was somewhat more complicated. Citing examples like Dickens' *She was busy preparing breakfast* (David Copperfield 1897:425), the authors suggest that we are dealing with a development that is a complex interaction of endogeny and exogeny. The structural use of *(busy* + V + *-ing)* is a continuation of an English pattern, with the main (and significant) influence from Afrikaans being a lifting of the restriction of *busy* to collocation with activity verbs.

Lass and Wright did not provide any examples from early South African English in support of their position. It is now possible to confirm via two databases of early South African English that their conclusions are essentially correct. The first corpus (examined by Mesthrie and West 1995) comprises the extensive letters written by first generation British settlers in the eastern Cape to the governor in Cape Town. Although our search was neither complete nor computer-assisted, we feel safe in concluding that *busy* in its present form was rarely used by this group (in fact we failed to find a single attestation in the numerous letters we perused).[3] However, we did not feel confident enough to conclude that *busy* was categorically absent in Settler speech. This position seems vindicated by the second corpus of interest to this study: the list of citations in the published version of DOSAEHP as well as the more numerous citations that remained unpublished. The earliest citation in this corpus is from the travels of George Thompson (1827:92):

(1) When we had outspanned at mid-day, and were busy cooking a mutton chop for dinner, we were startled by the appearance of two lions...

To assist in our description and analysis, we will use 'BUSY' (in small caps and single inverted commas) to denote a semantic reading of the word *busy* that coincides with the Concise Oxford Dictionary's sense of 'occupied, working, engaged, with attention concentrated' and 'NON-BUSY' for a semantic reading that largely excludes these meanings, especially the 'working' and 'attention concentrated' part. Sentence 1 above seems to refer to 'BUSY' activity. Another example from the same text (Thompson 1827: 98) shows an activity that seems ambiguous between 'BUSY' and 'NON-BUSY' activity.

(2) The part pierced by the arrows he had cut out and thrown away; the rest of the carcase he and his family had carried to their hut, and were busy feasting on it.

---

[3] Because of the age and state of the materials photocopying is not allowed; the corpus was too vast to copy out manually or on computer, given our lack of resources.

The next citations are from the published journal of Rev. Appleyard (1971), a missionary with the Wesleyan Society:

(3) They had just slaughtered a beast at the great place and were busy preparing for eating. (8 May 1841).

That this denotes 'BUSY' activity can be seen in the subsequent line of the journal:

(4) The breast, according to custom, was allotted to the chief and his great men, whom we found sitting under the fence of the calf kraal, whilst all his men were busy in the cattle kraal.

Appleyard is thus a traditional user of *busy* with activity verbs; indeed we find a more traditional example (now archaic) in his journal, showing *busy* in a locative sense:

(5) I have been very busy since the District [meeting] in compiling a Kaffer grammer (sic) according to my own views. (27 June 1842).

The only other nineteenth-century citations we find in the DOSAEHP files are as follows:

(6) One good thing we have, i.e. a coal fire; Jim is busy roasting his coffee over it. (Roche, *On Trek in the Transvaal*, 1878:321).

(7) The people are busy considering the question of union with the Transvaal Republic under President Pretorius (*Cape Monthly Magazine*, 21 Jan 1859, p12).

These nineteenth-century examples allow us the following conclusions, which are consonant with Lass and Wright's predictions. *Busy* was part of the early input into SAE (the language of visitors, missionaries and settlers); it mainly denoted 'BUSY' activities; but there was an option in the system for it to be stretched to cover activities that could only in a loose sense be construed as 'BUSY'. By contrast, the illustrations from the twentieth century show an explosion of examples of *busy* with 'NON-BUSY' verbs. Lass and Wright (1986:214-219) attribute the semantics/pragmatics, but not the syntax of the construction to influence from Afrikaans. However, even the concession to the semantic influence of Afrikaans might be premature. It is true that Afrikaans speakers sometimes feel compelled to translate (or 'transfer') the *besig om te* construction when speaking English. However, this frequently results in the non-standard construction (even for South Africa) *busy to V*. Such interlanguage forms sometimes turn up in the competent L2 writing of Afrikaans speakers, as in *A new programme was busy to develop* (= 'A new programme was busy developing/being developed'). An early example of this sort comes from the pen of an Afrikaans speaking bureaucrat working for the newly established British administration in the Cape in 1815. In a document with an otherwise more-or-less standard English grammar, the sentence given as (8) below occurs:

(8) ...they are very busy to collect a whole body of Kaffers to assist them in their project. (J.F. van de Graaff, letter to Major Fraser, 10 Nov 1815)

However tempting it is to see such a sentence as a forerunner to semi-auxiliary *busy*, there is a need for caution for three reasons: firstly, a sentence like (8) has *busy* as an adjective (qualified by the intensifier *very)*, rather than as a semi-auxiliary. Secondly, the verb form is the infinitive, not the participial of L1 SAE. The third reason has to do with comparative dialectology, rather than the structure of the construction. Sentences with semi-auxiliary are increasingly turning up in U.S. English, in genres that are marked as satirical or humorous:

(9) Pointed threats, they bluff with scorn
Suicide remarks are torn
from the fool's gold mouthpiece
the hollow horn plays wasted words
proves to warn
That he not busy being born
Is busy dying
(R. Zimmerman, *It's All Right Ma, I'm Only Bleeding)*

(10) She's busy fighting with my dad now. Can she call you back later? (*Dennis the Menace* cartoon, H. Ketcham)

(11) While you were busy being convinced it turns out we were late for class. (*Jughead* Comic, F. Doyle).

Sentence (12) taken from a newspaper in Berkeley is, however, devoid of a stylistically marked context:

(12) This industrious pumpkin spider set up camp near Chancellor Berhdahl's steps at University House. At this time of year, many of its relatives are busy emerging from pumpkin patches, which is where they got their name. (Berkeleyan, Oct 28-Nov 3, 1998).

In sentence (12), the collocation of *busy* with *emerging* marks a 'NON-BUSY' activity in a rather South African way. This would appear to weaken the case for Afrikaans influence over the SAE construction. In future work I intend to make more detailed comparisons based on corpora from various parts of the English speaking world, in order to ascertain whether we are dealing with a historical continuation or independent developments in the U.S. and South Africa.

3 A Grammatical Analysis

In the rest of this paper I offer a preliminary synchronic analysis of *busy* in SAE, with particular focus on the lexical and grammatical contexts in which it is already in use. This will feed into a brief discussion of its diachronic paths of grammaticalisation. The analysis will cover the range of verb types, semantic features and case roles that feature in the *busy* environment. As far as possible I will limit myself to actually attested examples from a variety of oral and written sources in SAE. But as is customary in modern syntax, it will also be necessary to stretch the analysis further by devising hypothetical test cases. These will be indicated where they occur; the majority of sentences in this paper are actually attested.

3.1 Stative versus Dynamic Verbs

The distinction between stative and dynamic verbs is an obvious candidate to use in refining our understanding of the use of *busy*. We might expect that dynamic verbs (defined as those that take *be + -ing* progressives and admit of imperative forms - e.g. *She was eating cake*; *Eat cake!*) easily collocate with *busy*; while stative verbs (defined as those that don't admit *be + -ing* progressives or imperatives - e.g. **She was understanding Greek*; **Understand Greek!*) should disallow collocations with *busy*. Whilst this is true as a first approximation, there are some finer details that can be teased out by paying closer attention to sub-types within each broad class and to differences in dialect use within SAE.

A semantic classification of verbs that I find particularly helpful here is that of Quirk, Greenbaum, Leech and Svartvik (1972:93-97). They divide dynamic verbs into the following classes:

> Activity verbs: (*abandon, ask, beg, call, drink*, etc).
> Process verbs: (*change, deteriorate, grow, mature, slow down* etc.)
> Verbs of bodily sensation: (*ache, feel, hurt, itch*, etc.)
> Transitional event verbs: (*arrive, die, fall, land, leave* etc.)
> Momentary verbs: (*hit, jump, kick, knock, nod*, etc.)

The list involves prototypical meanings and uses of the verbs; one of the problems associated with such a classification is that verbs may be used in more than one way; e.g. *change* as a (intransitive) process verb or (transitive) activity verb. By the norms of the definition in the *Concise Oxford Dictionary* for formal standard English only activity verbs allow *busy + be + -ing* constructions. This is not true of SAE where four of the categories collocate with *busy* without any problem for most speakers. Thus one may speak of people (or things) *busy drinking* (activity); *busy deteriorating* (process); *busy landing* (transitional); and *busy jumping* (momentary). These will become apparent in the course of our illustrations in the rest of the paper. The one doubtful case concerns verbs of bodily sensation. I have no data to prove or disprove the point, but my first impression is that 'BUSY-NESS' stops here. Sentences like the hypothetical (13)-(15) admit *(be + -ing)*, but not - I think - *(busy + be + -ing)*, even in SAE. However, in the light of data that might seem even more strange (see e.g. 26 and 29 below), but which are in fact attested in SAE, I will use the syntactician's category of judgement of 'questionable' rather than 'outright unacceptable'.

(13) */? I'm busy feeling sick.

(14) */? My tummy is busy aching.

(15) */? My nose is busy itching.

I propose the following hierarchy for the semantics of verb that collocate with *busy* verbs in SAE:

> *ACTIVITY > MOMENTARY > TRANSITIONAL EVENT > PROCESS > BODILY SENSATION.*

If verified by a corpus count, this hierarchy would also be a useful heuristic for cross-dialectal comparisons of the semantics of *busy*. Formal, standard British English stops at the first level; SAE goes all the way. It would be a matter of no great surprise to learn of other varieties falling somewhere in between (including informal varieties of British English).

To come to stative verbs, SAE disallows any collocation of these with *busy*. This applies to both sub-types of statives verbs defined by Quirk et al (1972:96): verbs of perception/cognition (e.g. *understand, love*) and relational verbs (*belong, concern*). This may seem self-evident given that the semantics of *busy* in SAE is that of 'in the process of', whilst stative verbs exclude processes from their denotation. The low position on the hierarchy for verbs of bodily sensation is unsurprising because, although they allow *be + -ing*, they are on the semantic borderline between stative and dynamic: c/f *I'm feeling sick* versus *I feel sick*.

There is one variety of English in South Africa that does allow stative verbs to take *be + -ing* endings. The following sentences are attested in Black South African English:

(16) I am knowing about it for a long time.

(17) He is having a cold.

(18) I am understanding it now.

I suspect that equivalents with *busy are* ungrammatical even in black South African English.

(19) *I am busy knowing about it for a long time.

(20) *He is busy having a cold.

(21) *I am busy understanding it now.

If this is so, then for the sociolect in question the distinction between stative and non-stative shows up better by collocations with *busy*, rather than *be + -ing*.

3.2 Subject Types

In formal, standard British English *busy* takes [+human] subjects only, though one can imagine metaphoric contexts in which this restriction is relaxed, with animals and perhaps modern machines acting in 'BUSY' human ways. In contrast, although [+ HUMAN] in SAE comes at the top of a hierarchy of the semantic features of 'BUSY' subjects, other features like [+ANIMAL], [+MACHINE], [+ (OTHER) INANIMATE] and [+DUMMY SUBJECT] are not excluded in 'ordinary' speech. As this will become evident in the examples in the ensuing section, I will not illustrate this claim here.

3.3 Semantic Roles

A good way of analysing the diffusion of *busy* in SAE is via a study of the semantic roles played by its subjects. Of course semantic roles are not independent of the notion 'type of verb', as categorised above, though we shall see that they do add to our understanding of *busy*.

3.3.1 AGENT

This is the expected role of NP subjects that take *busy* in international English. Sentences like the following form the vast majority in my corpus, showing AGENT to be the least marked case:

(22) Two young men were busy kicking in the display windows of a large departmental store when an AA patrolman appeared on the scene. (*Weekend Argus*, 1 April 1978.

(23) He was busy calling the union boys a hypocritical elite and a lot of other names. (N. Gordimer, *A Guest of Honour*, 1971:316).

(24) 584 Merino ewes with lambs (or busy lambing) (*Daily Dispatch* 1 December, 1986).

It should be clear that in SAE *busy* shows generalisation from 'BUSY' AGENTS to AGENTS in the process of doing things. It may also occur quite naturally, though less frequently, with other semantic roles.

3.3.2 PATIENT

Busy allows PATIENT subjects in active as well as passive sentences.

(25) You're busy losing weight! (Small talk with barman by University of Cape Town professor, 1998)

(26) The woman was very bad indeed, wrecked and racked on the golden vine, sodden beyond sensibility and busy dying. S.A. Botha, *Frontline*, 10 February 1990. (DOSAEHP 1996:129)

The collocation *busy dying*, which has been described as bizarre by an overseas colleague, is not unique to sentence (26). In late 1993 there was a news broadcast on national television in which a policeman describing a gruesome murder said, in a state of agitation, that by the time he got to the scene the victim *was busy dying* (noted independently by Roger Lass and myself).

The PATIENT role may also occur with passive sentences:

(27) Brand new home, busy being completed, comprising three bedrooms, lounge, diningroom... (*Evening Post* 28 July 1977).

(28) At that time my children were busy being born. (L2 Cape Flats English speaker to R.M.)

(29) The Cohens were busy being divorced in court. (*Best Jewish Jokes* 1968:11)[4]

PATIENTS in ergative constructions also occur, though they are rare:

(30) The plane is busy boarding now. (L2 Black English speaker)

Sentences 31 and 32 are hypothetical ergative sentences that I suggest are acceptable in SAE:

(31) The ball was busy rolling down the hill.

(32) The potatoes are busy cooking.

3.3.3 Other Semantic Roles

It is surprising how many of Fillmore's cases (1968) do not idiomatically co-occur with progressive aspect. Only at a pinch, given a suitable context, would one accept progressive versions of some of Fillmore's famous sentences:

(33) The key was opening the door. (Instrument)

(34) The hammer was breaking the window. (Instrument)

(35) The wind was opening the door. (Force)

---

[4] This joke book was published in the U.K., though it is clear that there is a strong South African connection involved. Streets of Johannesburg as well as South African currency (rands) are mentioned without contextualisation.

(36) The snow was ruining the garden. (Force)

How do these roles interact with *busy*? I do not have any examples of INSTRUMENT in my corpus; but presumably a context that allows any of (33) and (34) would allow an SAE colloquial equivalent with *busy*. One example with FORCE occurs in the DOSAEHP collection:

(37) It is busy snowing here at the moment (postcard from South African visitor in London).

Presumably, it is not just dummy FORCE subjects that behave this way; the hypothetical example (38) below seems possible in SAE:

(38) The rain is busy falling on the hilltops.

More generally, hypothetical sentence (39) with a 'FORCE' subject seems possible in SAE, given an appropriate context:

(39) The wind was busy opening and shutting the door.

Like FORCE and INSTRUMENT, the EXPERIENCER role (associated with cognitive or psychological verbs) does not appear to favour progressive *be* + *-ing* in standard English.

(40) John heard an explosion.

(41) *John was hearing an explosion

(42) Jill was angry.

(43) Jill was being angry (grammatical as AGENT putting on an act, not as EXPERIENCER).

No attestations of *busy* with EXPERIENCER roles occurs in my corpus. People are said to be *busy* listening to the news (AGENT), but not, to my knowledge, to be * *busy hearing the news* (EXPERIENCER).

Sentence (44) shows the ESSIVE role in SAE (associated with being and becoming).

(44) He slows down only while the robot is *busy* turning red.

I paraphrase the sentence to mean the robot (= 'traffic light') was in the process of becoming red. A hierarchy for case roles that favour *busy* can be tentatively posited:

> AGENT > PATIENT > FORCE (dummy subjects) > INSTRUMENT > ESSSIVE

4 Further Grammaticalisation in SAE

We might say that *busy* in SAE is busy being grammaticalised. In the history of English it was originally used as an adjective followed by the preposition *in* plus a verbal noun in *-ing*. The preposition was dropped, and the resulting construction reinterpreted as *busy* (adj.) plus verb + participle. Thus Jane Austen could write (45), where 'busy' is used ironically. Nevertheless the status of *busy* is that of adjective, as evidenced by the intensifier *very*.

(45) Our whole Neighbourhood is at present very busy grieving over poor Mrs Martin, who has totally failed in her business, and had very lately an execution in her house. (Letter from Jane austen to Cassandra, 25 October 1800, cited by Lass and Wright 1986:217)

In SAE *busy* it is being bleached of the 'BUSY' meaning and turned into an auxiliary verb denoting ongoing activity plus verb + participle. This is evident in its prosody; it is unstressed, and disallows the pause that is characteristic of a phrase like *busy, working in the garden* in older standard British English. Furthermore one cannot say idiomatically the hypothetical sentence 40 in SAE:

(46) *It's very busy snowing here at the moment.

As a grammatical morpheme indicating progressive activity rather than a content morpheme indicating necessarily 'BUSY' activity, it collocates with almost any subject that allows ordinary progressive aspect. In the mid- to late 1990s several interesting sentences are turning up showing further degrees of grammaticalisation, and an almost complete bleaching of *busy* in its semi-auxiliary contexts. This is evident in sentences (47)-(50):

(47) RM: I think he's quite young?
S.B. Yes, I think he's still busy studying. (= 'Yes, I think he's still a student')

(48) I am currently busy compiling the Arts Orientation Guide (University of Cape Town, Arts Students' Council letter, 1996).

(49) Our flight attendant is busy getting permission for take off at the moment. (South African Airways steward, 1998).

(50) I'm busy becoming acquainted with my data (Graduate student to R.M. 1998)

The use of adverbials (*still, currently, at the moment*) in (47)-(49) and the collocation of busy with the process verb *becoming* in (50) suggest that not only has *busy* been completely bleached of its 'BUSY' semantics, but that speakers feel the need to supply adverbial material stressing the present progressive nature of their utterance. A further degree of grammaticalisation is suggested by the exchange in (51):

(51) R.M.: Is it [= a piece of electrical cord] frayed?
S.: Yah, it's busy.

What S. intended (as a repetition of my question showed) was 'Yes, it's busy being frayed'. This participation in an elliptical construction with be, shows that to a large extent, *busy* + *be* + *-ing* is equivalent to *be* + *-ing*.

5 Conclusion

There are many loose ends which still need to be tied up in this preliminary study. A broader data base involving international corpora is being constructed. Closer attention also needs to be paid to social and stylistic factors: sub-groupings like L1 or L2 speakers; hyper-casual or relatively neutral style, written or oral speech etc.

References

Branford, Jean (with William). 1991. *A Dictionary of South African English*. Cape Town: Oxford University Press. (4th ed.)

Branford, William, Hacksley, Malcolm and Silva, Penny (eds.) 1996. *A Dictionary of South African English on Historical Principles*. Oxford: Oxford University Press.

Fillmore, Charles 1968. 'The case for case'. In Emmon Bach and Robert Harms (eds) *Universals in Linguistic Theory*. New York: Holt, Rhinehart and Winston.

Frye, John (ed.) 1971. *The War of the Axe and the Xosa Bible - the Journal of the Rev. J.W. Appleyard.* Cape Town: Struik.

Lass, Roger and Wright Susan 1986. 'Endogeny versus contact: "Afrikaans influence" in South African English'. *English World Wide,*7,2: 201-223.

Quirk, Rando!ph, Greenbaum, Sidney, Leech, Geoffrey and Svartvik, Jan 1972. *A Grammar of Contemporary English.* London: Longman.

Department of Linguistics and Southern African Languages
University of Cape Town
Private Bag Rondebosch
7700
raj@beattie.uct.ac.za

The Emergence of Creole Subject-Verb Agreement

Miriam Meyerhoff

This paper undertakes two tasks: first, it resolves an outstanding question as to the proper structural description of Bislama; second, it will unify this analysis of the verb system with the distribution of phonetically null subjects in Bislama. In this way, it will show that resolving patterns of variation in one domain of the grammar contributes to a meaningful account of what underlies variation in another domain.[1]

We first turn to the data. Table 1 shows the paradigm for finite verbs in Bislama (the English-lexified creole spoken by c.180,000 people in Vanuatu, SW Pacific) in all persons and numbers with the verb *karem* 'to carry; to bring'.

	SINGULAR	DUAL	TRIAL	PLURAL
1 (incl.)	--	*yumitu* karem	*yumitri* karem	*yumi* karem
1 (excl.)	*mi* karem	mitufala **i** karem	mitrifala **i** karem	mifala **i** karem
2	*yu* karem	yutufala **i** karem	yutrifala **i** karem	yufala **i** karem
3	hem **i** karem	tufala **i** karem	trifala **i** karem	olgeta **oli** karem

Table 1: Paradigm for finite verbs in Bislama, *karem* 'carry, bring'.

Like many of the substrate languages that continue to be spoken in Vanuatu, Bislama makes four distinctions in number, and in the first person non-singular forms also distinguishes a 'we' that is inclusive of the addressee, and a 'we' that excludes the addressee. In each cell, *karem* is preceded by first the pronoun and then a morpheme most commonly realized as *i*. At two points in the paradigm the pattern of *i* marking breaks down: one is in the third person plural, where the morpheme *oli* intervenes between the verb and the pronoun. Like *i*, it is clear that *oli* is cliticized to the main verb. Example (1) shows that although there is some flexibility in the placement of adverbials, exemplified in (1b) by the adverb *bae* that indicates irrealis mood, (2)

---

[1] This paper revises material from chapters 5 and 6 of my 1997 dissertation. I am grateful to the Wenner-Gren Foundation for their support (grant #5742) and to Sharon Morrie Tabi in Santo. I also thank the following individuals for comments on earlier versions: Terry Crowley, Gillian Sankoff, Bill Labov, Naomi Nagy. For discussions on split *pro*-drop systems I thank David Heap and Richard Kayne. None of the above necessarily agree with the analysis presented here.

BEST COPY AVAILAB

shows that it is ungrammatical when placed between the verb and *oli*. The same facts hold for *i*, and for other adverbs, e.g. those with focus functions.

(1) a. *Ol pikinini oli krae from kek*
The children cried out for cake. (M-94-3, Lolan)
b. *(Bae) ol pikinini (bae) oli krae from kek.*
The children will/would cry out for cake.

(2) **Ol pikinini oli (bae) krae from kek.*

The second place the paradigm in Table 1 breaks down is the first and second person singular and the first person inclusive non-singular forms (highlighted in italics). We turn now to a consideration of what the proper analysis of these cells is.

1 Prior Accounts of the Data

Prior accounts of Bislama have analyzed these forms in three ways. Tryon's (1987) analysis is the most economical. Referring to these morphemes in his pedagogic grammar as 'predicate markers' (a rather meaningless descriptor, which nevertheless is widely used in discussing Bislama and the related language Tok Pisin), Tryon analyzes Bislama as having a predicate marker, *oli* in 3p, a zero predicate marker in 1s, 2s and 1incl. forms, and the predicate marker *i* elsewhere.

Crowley (1990), analyzes the system as having an underlying *i* in all persons and numbers (except 3p). He suggests that the *i* is assimilated to a preceding high vowel. That is, Crowley argues that (3b) is the underlying form for the surface form (3a).

(3) a. *Mi harem wan noes long naet*
I heard a noise in the night
b. *Mi i harem wan noes long naet (/mi + i / > /mi/)*

This analysis would seem to be much influenced by the (to my knowledge, undisputed) analysis for the Tok Pisin. Tok Pisin, while closely related to Bislama, differs in some crucial ways from Bislama in the finite verb system. Tok Pisin has *i* marking in careful speech with 3p, and wide-spread deletion of *i* marking in all persons and numbers in casual speech. Neither of these properties hold for Bislama. Moreover, Crowley's analysis is somewhat at odds with some of his other data. For instance, in (4a) he records that when adverbial modification of a 1s verb does not result in *i* surfacing, instead we find doubling of the pronoun. As (4b) shows, this results in a pattern is not analogous to pronoun doubling in other persons, rather it seems to suggest that the second *mi* in (4a) is functioning as the *i* in (4b) does.

(4) a. *Mi bae mi no kam*
I won't come. (Crowley 1990:235)
b. *Hem bae i no kam/ *Hem bae hem i no kam.*
S/he won't come.

Finally, Guy (1974), Camden (1977) and Charpentier (1979) independently forward a radically different analysis. They note the comparative frequency of utterances with the form of (5). They also note strong structural parallels between the Bislama verb phrase and the patterns found in many of the substrate languages in Vanuatu. On the basis of such parallels, they argue, what appears (on the basis of its function in the lexifier) to be a pronoun in the cells for 1s, 2s and 1 incl. in Table 1 is no pronoun at all. It is an agreement marker, or, in Charpentier's terms "des modalités personelles" (1979:307). Utterances in which there is a singleton *mi* are, they argue, cases of *pro*-drop leaving only an agreement marker that happens to be homophonous with the pronoun and the main verb.

(5) *Mi mi kakae*
1s 1s eat 'I ate/eat'

In sum, the difference between the three analyses is this: Tryon and Crowley believe *mi* and *yu* in Table 1 function only as subject pronouns in Bislama; Guy, Camden, and Charpentier (GCC) believe *mi* and *yu* function as both subject pronoun and as subject-verb agreement.

Clearly, these are fundamentally different analyses. In addition, they fail to adequately define what *i* and *oli* actually are in Bislama, resorting to terms like 'predicate marker' or 'marker of person modality'. In the next two sections, I will argue that Tryon and GCC are both right. The morphemes *i* and *oli* are bona fide subject-verb agreement, as is found in other Vanuatu languages, but the agreement in 1s and 2s is a zero morpheme.

I will make my case using quantitative data taken from a corpus of conversational Bislama recorded in northern Vanuatu in 1994–5. Forty-two speakers, ranging in age from 8 to approximately 65 years were recorded in their own homes or when stopping to visit with the researcher. In all cases, day-to-day conversations with friends, family and the researcher were recorded, though two of the children under 12 years were recorded telling stories. Tapes were made both in the northern township, Santo, and in a village on a nearby island. Roughly equal numbers of women and men were recorded. Segments of conversation were transcribed to generate a corpus of c. 30,000 words; all examples that follow are drawn from the corpus, unless noted otherwise. Speakers are identified with pseudonyms.

2 The Proper Description of the Bislama Verb Phrase

In this section, I review quantitative evidence for and against the three analyses outlined above. It should be noted that the analysis is based primarily on data on 1s and 2s forms. There were very few of the 1incl. forms in the corpus examined. Since Crowley's and Tryon's analyses differ only in the detail (in principle agreeing that *mi* and *yu* are simply pronouns), I will evaluate first the evidence for each of these positions, ultimately showing that the weight of the evidence is in favor of Tryon's proposal. Having determined which of the two pronominal analyses is the more appropriate, I will evaluate how well the pronominal and subject-verb agreement analyses of *mi* and *yu* fit the data.

2.1 Evaluating the "Pronoun Only" Analyses

In evaluating Tryon's and Crowley's analyses, we will consider all evidence that might indicate there is indeed an underlying *i* in 1s and 2s. We will examine the strength of the evidence in three domains that would support Crowley's analysis of *i* assimilation to a preceding high vowel.

First, let us note that it is clear Bislama allows phonetically null subjects (since this is true in both main and subordinate clauses, and there is also no overt expletive subject, we can refer to Bislama using the short-hand, calling it a *pro*-drop language). However when the subject is interpreted as 1s (or 2s) but is phonetically null, no underlying *i* morpheme resurfaces. Since Crowley made no explicit claims for the ordering of the *i* assimilation rule in Bislama, it would be possible for his analysis of *i* assimilation to hold if its application were ordered after some rule of subject deletion applied.

But Crowley's analysis also fails to square well with two other facts. Let us consider in more detail the facts alluded to in (4a) and (4b). The data shows that adverbials modifying the verb phrase (*bae*) or the subject (*wan nomo*) can intervene between a 3s subject and the complex *i* + *V*, without any change to the unmarked inventory and order of morphemes in the clauses. However, this is not the case when the subject is 1s or 2s.

This data is extremely problematic for Crowley's analysis. Not only does the adverb fail the make the underlying *i* he posits appear, but even if the assimilation rule were triggered before adverb insertion (an ordering that would be unusual in itself), an apparent copy of the pronoun is obligatory in these cases. This evidence favors the GCC analysis, which we return to shortly.

The second fact that needs to be considered is that Crowley's analysis predicts that whenever a subject ends in a high vowel in other persons and numbers, the *i* morpheme should also assimilate. In fact, an analysis of such tokens in the corpus tends not to support this. All complex NPs where the final segment is a high vowel (6-7), and all proper names ending in high vowels (8) in the corpus were examined.

(6) a. *Afta mi pusum [hed blong mi] i go*
So then I pushed my head through. (M-95-19, Mesek)
b. **Afta mi pusum [hed blong mi] go*
(7) a. *[Woman blong yu] i kam long wea?*
Where does your wife come from? (S-94-3, Simeon)
b. **[Woman blong yu] kam long wea?*
(8) a. *(Tammy/Wili/Lili) i no save*
T/W/L didn't realize it. (S-95-11, Juliet)
b. *#Tammy no save*

Table 2 summarizes the quantitative results. It shows that an overt *i* almost categorically follows complex NP subjects ending in a high vowel. The data with proper names ending in high vowels is mixed, showing a slight tendency for assimilation.[2]

Subjects w. /i/ or /u/ final	Form of predicate		**Total**
	i + V	Ø + V	
(10) NP + *blong mi*	9	3	12
(11) NP + *blong yu*	6	0	6
(12) Proper Name	12	14	26
Total	27	17	44

Table 2: Presence and absence of *i* when immediately preceded by a high vowel (non-subject *mi* and *yu* and Proper Names).

---

[2] The presence or absence of *i* in these contexts was verified by Sharon Tabi, a Bislama-dominant speaker from Pentecost.

In sum, although the data does not uniformly line up in favor of one analysis over another, the bulk of the evidence suggests that if the morphemes in question are best analyzed as pronouns, then Tryon's analysis of a zero 'predicate marker' in 1s and 2s fits the data better than Crowley's analysis of an underlying *i*.

2.2 Evaluating the 'Subject Agreement' Analysis

Just as the competition between Crowley's and Tryon's analyses had to be resolved by turning to evidence elsewhere in the grammar (phonology-syntax interface), weighing the goodness of fit of Tryon's analysis and the GCC analysis requires examining data from another domain of the grammar (discourse-syntax interface). In this section I will compare the distribution of 3s pronouns and *i* with the disputed forms *mi* and *yu* in a range of discourse conditions.

The hypothesis is as follows: If *mi* and *yu* only function as pronouns in Bislama (Tryon's contention), then we expect to find them distributed in different discourse contexts with the same relative frequency as we find other, indubitable, pronouns such as the 3s *hem*. We should also find, *mutatis mutandis*, that the frequency of phonetically null 3s subjects (with the form *Ø i V*) is much the same as bare verbs where the interpretive subject is 1s/2s (the form *Ø V*). And we should also expect to find that the frequency of focused 3s subjects (with the form NP_i, hem_i *i V*) in each discourse context is much the same as the frequency of forms like *mi mi V*.

$$\frac{mi\ mi\ V}{NP,\ hem\ i\ V}\ \text{VS}\ \frac{mi\ V}{hem\ i\ V}\ \text{AND}\ \frac{\emptyset\ V}{\emptyset\ i\ V}$$

Figure 1: Within group similarities expected in the distribution of 3s and 1s/2s subjects if *mi* and *yu* are pronouns.

However, if the Guy-Camden-Charpentier analysis is correct and *mi* and *yu* may be subject-verb agreement markers homophonous with the pronouns, then we expect to find their distribution across different discourse contexts to be different. Specifically, we expect that a singleton occurrence of *mi* will be distributed in a manner comparable to 3s forms where the pronoun has been dropped and only the agreement marker *i* remains. Utterances in which there is a sequence of *mi mi* + *V* should be found with comparable frequency in different discourse contexts as focused 3s subjects, i.e. NP_i, hem_i *i V*.

$$\frac{mi\ \mathbf{mi}\ V}{hem\ \mathbf{i}\ V} \quad \text{vs} \quad \frac{\mathbf{mi}\ V}{\emptyset\ \mathbf{i}\ V}$$

Figure 2: within group similarities expected in the distribution of 3s and 1s/2s subjects if *mi* and *yu* are pronouns *and* (homophonous) subject-verb agreement.

It is important to note that the prediction is not that the relative frequency of these forms will be the same *between* groups. For example, the prediction is not that 1s/2s subjects will be phonetically null as often overall, or even as often in a specific interclausal context as 3s subjects are. The similarity that is required is a *within* groups similarity. This means that if an interclausal discourse context favors phonetically null 3s subjects more than another, it is predicted that the same contexts will be ranked in the same order relative to one another when the subject is 1s/2s.

Five interclausal relations were distinguished based on the form and the grammatical role of the referent. These were cases where:

A	the subject of the (current) clause had also been the subject[3] of the preceding clause;
P	the subject of the (current) clause was some other argument in the preceding clause; and
N	the subject of the (current) clause was not present in the preceding clause.

where 'clause' was defined as a finite verb, whether main or subordinate. Non-finite clauses (including imperatives) and the second verb in serial verbs expressing motion or location, both of which never have an overt subject, were excluded from the data. In the first two conditions, the antecedent referent was further discriminated on the basis of its form, overtly realized vs. phonetically null (giving *A-o* 'overt subject', *A* 'phonetically null subject', *P-o* 'overt other argument', and *P* 'phonetically null other argument'). There is no *a priori* reason to suspect that further constraints on the focusing of subjects apply when the referent is the speaker or addressee but do not apply when the subject is a third party, thus the experiment seems valid.

An analysis of this data using Goldvarb 2.1 (Sankoff et al. 1992) was conducted. Guy (1980) has shown that use of logistic regressions must be cautious when the data is highly skewed or very patchy, i.e. when the number of tokens in some cells falls below 30. In this data, the problem cells are clustered in the *P* and *P-o* conditions. A multivariate analysis of the data leaving out the data in these conditions was conducted but in each case this resulted in a model that was a significantly worse fit to the data than was obtained when the data from these conditions was included. Consequently, these factors are retained for the within groups comparison.

---

[3] Bislama has no passive, so this condition generally reflects a continuity in thematic role between clauses as well.

The results of the logistic regressions are represented graphically in Figures 3-7. The frequency of the different forms of subject in different discourse conditions is converted into probability weightings in the tables,

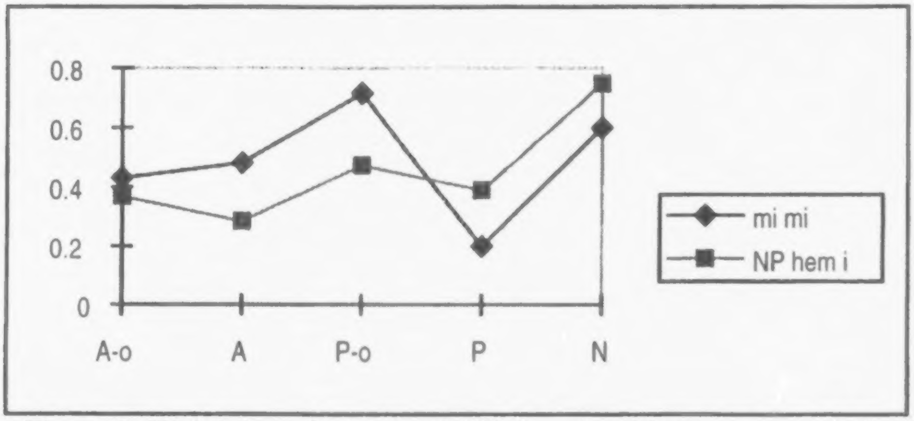


Figure 3: Weighting of *mi mi* and *NP, hem i* subjects compared (Analysis B predicts isomorphism). LINEST = 0.4908.

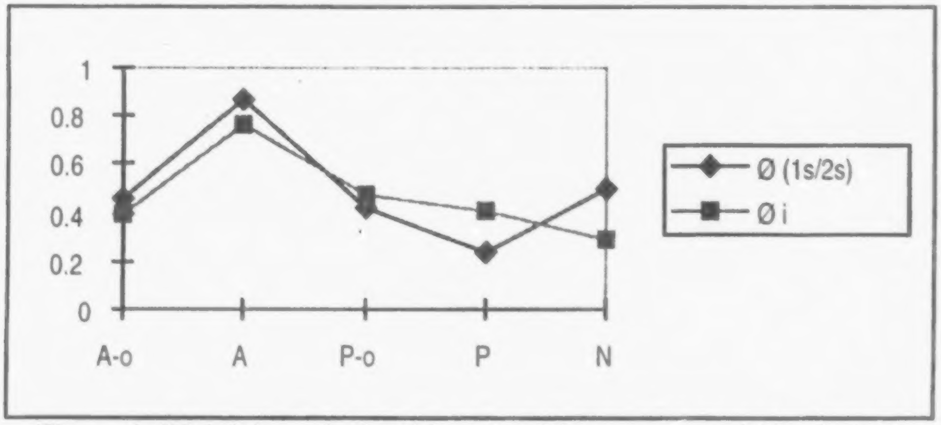


Figure 4: Weighting of *mi* and *hem i* subjects compared (Analysis B predicts isomorphism). LINEST = 0.4908.

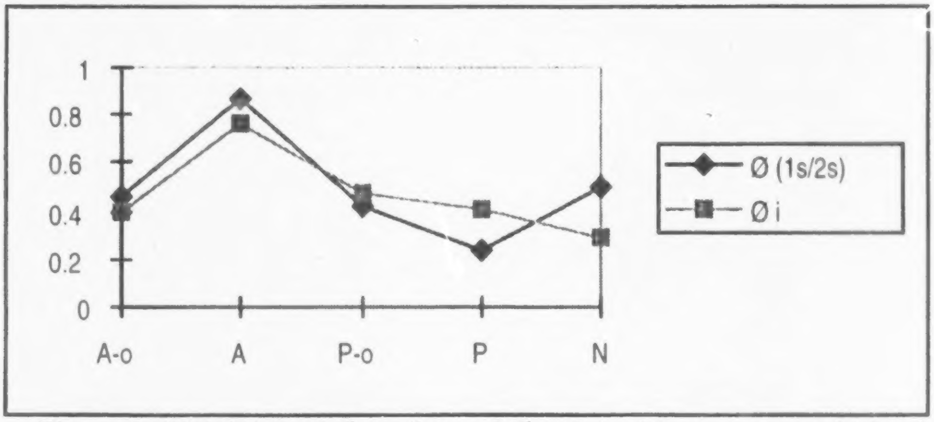


Figure 5: Weighting of *Ø (1s/2s)* and *Ø i (3s)* subjects compared (Analysis B predicts isomorphism). LINEST = 0.9711.

expressing the degree to which each condition favors or disfavors the different forms of subject. Figures 3-5 test the hypothesis that *mi* and *yu* are pronouns only; Figures 6-7 test the hypothesis that *mi* and *yu* are both pronouns and subject-verb agreement markers.

The degree to which there is comparable within group behavior of the different subject forms has been evaluated using the LINEST (linear estimate) function. This is a measure of how much the two paths being compared behave alike or diverge from each other (calculating closeness on the basis of the square difference between the two lines at each point of comparison). A value of positive one indicates that the two forms being compared are always moving the same way; a value of minus one indicates that they consistently move in different ways. Figures 3-5 show LINEST values that are consistently positive, and in the case of the hypothesized null subject condition (Figure 5) behavior of 3s and 1s/2s subjects across the different discourse contexts hardly differs at all. The modest difference in the way subjects with the form *mi mi V* pattern compared to the 3s subjects *NP, hem i V* suggests that we were correct to proceed assuming that there are no major independent constraints on how and when it is appropriate to focus a subject referring to the speaker or addressee compared to when it is appropriate to focus a 3s referent.

This is in marked contrast to the picture painted when testing the GCC analysis, as shown in Figures 6-7.

BEST COPY AVAILABLE

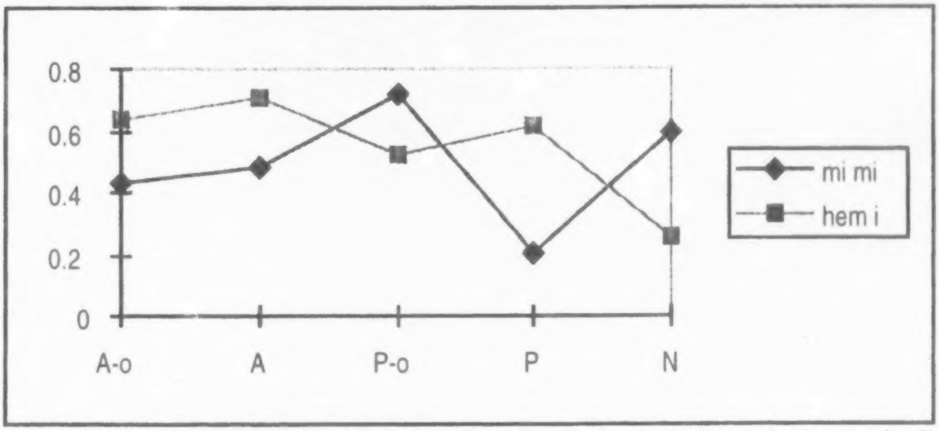


Figure 6: Weighting of *mi mi* and *hem i* subjects compared (Analysis C predicts isomorphism). LINEST = -0.491.

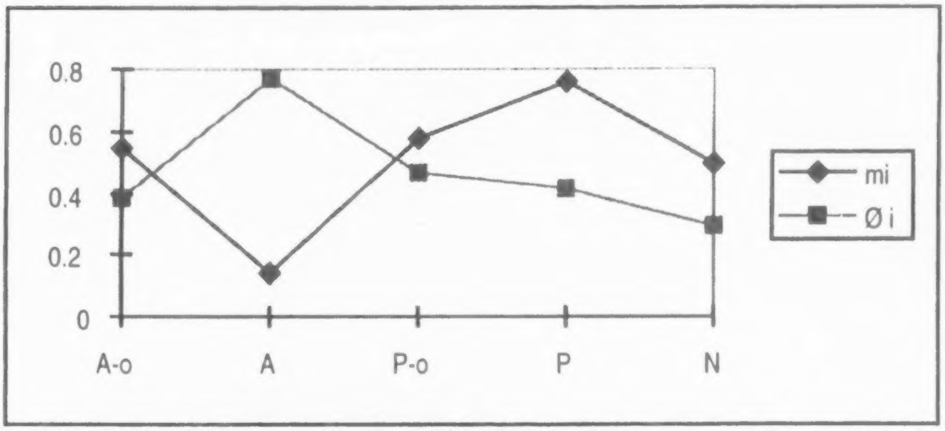


Figure 7: Weighting of *mi* and *Ø i* (3s) subjects compared (Analysis C predicts isomorphism). LINEST = -0.971.

Figure 6 shows that 1s/2s subjects with the form *mi mi* and *yu yu* are not distributed across the five interclausal environments in a manner comparable to the distribution of focused 3s subjects, while Figure 7 shows even more clearly that subjects with the form *mi* or *yu* alone do not pattern like 3s subjects that are phonetically null (leaving only the agreement marker *i*). We must conclude, then, that the singleton forms *mi* and *yu* are not subject agreement markers on the verb as GCC argue.

To sum up, then, on the basis of the distribution of different subject forms across a range of phonological and discourse environments, we conclude that *mi* and *yu* are not subject-verb agreement markers, and that the agreement marker analogous to 3s and 3p *oli* in these forms is a zero morpheme.

3 The Development of a Grammar Licensing *Pro*-drop

This claim will be supported by looking more closely at the patterns of phonetically null and overt pronominal subjects found in conversational Bislama. It will be argued that the distribution of these forms is best explained in terms of the morphosyntactic informativeness of the inflectional system outlined in section one. In doing so, we will see that the account of 1s/2s subject-verb agreement forwarded in this paper has more than mere descriptive adequacy. It provides the basis for a unified account of the distribution of phonetically null subjects as well.

There is a marked difference in the frequency of overt/null 1s/2s subjects and the frequency of overt/null 3s subjects. The relevant Ns are shown in Table 3:

sbj form	***mi & yu***	***hem***
Ø	118	1010
pronoun	1364	664

Table 3: Frequency of pronoun and phonetically null subjects in Bislama (1s & 2s compared with 3s).

Table 4 shows that this distinction between the persons generalizes across singular and plural. Table 4 shows the probability with which subjects in each person and number occur with a phonetically null subject (complete co-

Person and number of subject referent	**Goldvarb weighting**	**N clauses**
3s	0.709	1719
3p	0.869	846
1s	0.147	1054
1p	0.268	375
2s	0.206	397
2p	0.197	71

Table 4: Goldvarb weightings for all speakers for person and number of subject referent (0 = consistently pronoun subject; 1 = consistently null subject). Input probability for all speakers = 0.418

occurrence = 1) or a pronominal subject (complete co-occurrence = 0). The probabilities are shown as weightings derived by means of logistic regressions, again calculated using Goldvarb 2.1.

The results show that 3s and 3p subjects pattern alike, favoring a phonetically null subject, and that first and second person subjects singular and plural pattern alike, favoring the use of an overt pronoun in subject position. Unfortunately, the historical record on the use and distribution of phonetically null subjects in Bislama is sketchy at best. Charpentier notes "In older pidgin [Bislama], the verb was always preceded by a noun or pronominal subject... Nowadays, the personal pronoun tends to be left out, leaving utterances that consist only of subject-verb agreement and the verb stem" (1979: 353).[4] This suggests that the possibility of omitting pronominal subjects is a grammatical innovation that emerged during the 1960 or 1970s, following the crystallization of the pronoun and subject-verb agreement system in Table 1 and during the period in which growth of the urban centers, especially Port Vila, in Vanuatu created a comparatively large pool of first language speakers of only Bislama.[5] Since these are precisely the conditions under which we would expect to see the effects of UG become stronger than effects of the substrate, the development of what appears to be a grammar of *pro*-drop is especially noteworthy. Moreover, the fact that Bislama appears to have grammaticized a split *pro*-drop system is of even greater significance.

In the pidgin stages of Bislama (mid-19th century), there was no agreement between subject and verb in any person and only a simple pronominal system (*mi, ae* '1s', *yu* '2s' and *i* '3s/p') was used. Crowley's data strongly suggests that by the turn of the century, the system had been reanalyzed so that 3s and 3p pronouns were distinct (*hem* '3s' and *olgeta(fala)* '3p'), with *i* being reassigned as an agreement marker (1990: 231, 243) . Since Crowley's data shows that even while *hem* was crystallizing as the 3s pronoun, *i* continued to occur in some clauses in what appears to be subject position (i.e. *hem* is absent). Whether these should be interpreted as holdovers from an earlier stage in the grammar (speakers using *i* as a pronoun), or whether they indicate the first examples of 3s null subjects cannot be determined. The fact of the variation is, however, vital.

---

[4] "Dans le pidgin ancien, le verbe était toujours précédé d'un nom ou d'un pronom sujet... Aujourd'hui, le pronom personnel tend à être abandonné, la modalité personnelle et le radicale verbal composant seuls un énoncé" (Charpentier 1979: 353).

[5] For example, the oldest speaker I recorded in Santo/Malo who spoke Bislama as a sole first language was born in 1972 in Port Vila.

Let us assume that the variation at this crucial stage created the possibility for speakers to interpret *i either* as a pronoun *or* as an agreement marker. This is not implausible given the immediate history of the forms (English *he* and pidgin *i* as subject pronouns) and the substrate models available (Oceanic languages with agreement prefixes on the verb). In other words, convergence between the lexifier and the substrate results in creole variation. This necessitates further analysis and generalization on the part of the language users. As a solid pattern for the third person agreement begins to take shape, a set of agreement morphemes emerge for first and second person as well. Maintaining the singular/plural contrast marked overtly in the third person, first and second distinguish singular and plural—respectively, Ø, and *i*. In short, we find that the data requires a synthetic account of creole variation and creole grammatical development like those proposed by Sankoff (1984) and Crowley (1990).

We lack the amount of quantitative data from the 20th century that would be needed to conclusively support this account, nevertheless some of the 1990s data suggests that this is the path along which Bislama continues to develop. The analysis of the entire conversational corpus failed to reveal any significant apparent time effects for this variable. However, a closer examination of a subset of the corpus showed that there does indeed seem to be some change taking place with respect to the variable of null subjects.

When the comparison was restricted to the children in the corpus and their immediate caregivers and their uncles and aunts, a very strong effect for age was found.[6] The children appear to be generalizing the patterns in the wider speech community, approaching a categorical treatment of first and second person subjects in particular.

In short, the data from the children and their caregivers suggests that the synchronic variation found in each person in Bislama is indicative of a diachronic change in progress, during which speakers are moving towards a less variable system, one that would be compatible with some of the structural analyses of split *pro*-drop systems observed in other language families.

One final question needs to be answered as to the probable outcome of this ongoing reanalysis. It is unlikely that the outcome of this will be a time in which speakers of Bislama categorically use null anaphors in the third person and pronouns in the first and second. We see a degree of variability remains even in the first and second person for the children, and we noted earlier that intersentential variables (referent's form and grammatical role) in the preceding clause play a part in constraining the distribution of null subjects in Bislama. An important consequence of their significance is that they remind us that speakers can bring numerous skills to bear in identifying the referent when the subject is phonetically null. While verbal inflections may be a preferred mode of identification, the fact that null subjects with first and second person referents are interpretable in Bislama (and more generally in languages with no verbal morphology to rely on) shows clearly that identification is by no means a unitary notion. The interaction between linguistic and subjective (or affective) variables will continue to be a factor, just as they are for prototypical null subject languages such as standard Italian. What I have shown here, however, is that they constrain a relatively small part of the variation observed in the form speakers use to express a subject in Bislama.

---

[6] Results of a regression analysis comparing children 12 and under with caregivers and young aunts and uncles, r^2(adjusted) = 71.6%, $p < 0.001$.

4 Conclusion

At the outset, the problem faced in this paper was a descriptive one: which of three accounts of Bislama finite clauses best fits the data? Evidence from phonology, syntax and discourse was considered and it was concluded that the preverbal clitics in Bislama are subject-verb agreement, and that they have three variants, the overt forms *i* and *oli* and a null variant (in 1s and 2s).

Next, the distribution of null subjects in Bislama was examined and it was seen that variation emerges in the grammar at precisely the point where the verbal morphology is maximally informative. As generative accounts of the null subject parameter would predict, null subjects are emerging as the norm in the third person but not the norm in first and second person in Bislama. This makes sense since third person subject-verb agreement is maximally informative in the third person, i.e., it satisfies the requirement that a null subject be identified.

References

Camden, Pastor Bill 1977. *A Descriptive Dictionary: Bislama to English*. Maropa Bookshop: Vila.

Charpentier, Jean-Michel 1979. *Le Pidgin Bislama(n) et le multilinguisme aux Nouvelles-Hébrides*. Language et Civilisations à Tradition Orale, 35. Paris: Sociéte d'Etudes Linguistiques et Anthropologiques de France [SELAF].

Crowley, Terry 1990. *Beach-la-Mar to Bislama: The emergence of a national language in Vanuatu*. Oxford: Oxford University Press.

Guy, Gregory R. 1980. Variation in the group and the individual: the case of final stop deletion. In William Labov (ed.) *Locating Language in Time and Space*. New York: Academic Press. 1-36.

Guy, J.B.M. 1974. *Handbook of Bichelamar/Manuel de Bichelamar*. Pacific Linguistics, C-34. Canberra: The Australian National University.

Sankoff, David, Pascale Rousseau, Don Hindle and Susan Pintzuk 1992. Goldvarb version 2.1. Adapted for the Macintosh by David Rand. Shareware, Université de Montréal.

Sankoff, Gillian 1984. Substrate and universals in the Tok Pisin verb phrase. In Deborah Schiffrin (ed.) *Meaning, Form, and Use in Context: Linguistic applications*. Washington DC: Georgetown University Press. 104-119.

Department of Linguistics
University of Hawai'i at Manoa
1890 East-West Rd
Honolulu, HI 96822
mhoff@hawaii.edu

The *University of Pennsylvania Working Papers in Linguistics (PWPL)* is an occasional series produced by the Penn Linguistics Club, the graduate student organization of the Linguistics Department of the University of Pennsylvania.

Publication in this volume does not preclude submission of papers elsewhere; copyright is retained by the authors of the individual papers.

Volumes of the Working Papers are available for $15, prepaid. Please see our web site for additional information.

The PWPL Series Editors

Jim Alexander
Alexis Dimitriadis
Na-Rae Han
Elsi Kaiser
Michelle Minnick Fox
Christine Moisset
Alexander Williams

How to reach the PWPL

U. Penn Working Papers in Linguistics
Department of Linguistics
619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104–6305

http://www.ling.upenn.edu/papers/pwpl.html
working-papers@ling.upenn.edu

Double Subject Marking in L2 Montreal French

Naomi Nagy and Hélène Blondeau

1 Using SLA to Analyze Grammatical Variation

Dennis Preston (1996:31) writes:

> SLA [Second language acquisition] is in some ways dramatically positioned, I believe, to contribute to variationist understandings of language. Its respondents are on a fast-track of language change, allowing real- rather than apparent-time studies.

This study focuses on a sample of second language (L2) speakers in order to identify which aspects of a grammar may be transmitted in a language contact situation and what social characteristics of the speakers promote or impede this transmission.[1] A second goal of this investigation is to describe the competence in French of the first generation of Montreal Anglophones (native English speakers) that had access to French immersion schooling and to understand which social factors determine their level of competence. We address the following three research questions, and determine that the answer to each is "yes."

(1) Does double marking, a pattern not taught in school, exist in L2 Montreal French?
(2) Does the rate of double marking correlate to the type and amount of acquisition and contact?
(3) Does the variation provide evidence of acquisition of L1 grammar (as opposed to mimicking of a salient surface structure)?

As members of a minority population in a city with a majority of French native speakers, young Anglophone adults have many types of contact with French speakers. Some grew up having regular interactions with French-speaking relatives and close friends; others had no French speakers in their social environment as young children but made French friends as adolescents. Still others had little or no social contact with Francophones until adulthood, but now have a Francophone spouse or partner. Another major factor is schooling. We include speakers with the three types of school background: ordinary French classes in English schools, immersion programs in English schools, and French schools. We examine the effects of these differences in background in analyzing the French that the Anglophones have acquired.[2]

---

[1] The authors gratefully acknowledge the support of a Summer Research Fellowship from the University of New Hampshire to the first author in 1997. We also thank P. Thibault and G. Sankoff for allowing us to use this corpus of Montreal L2 speech.

Prior studies of speakers from the same corpus have analyzed variable phonological patterns which distinguish Montreal French from "standard" European French (Nagy et al. 1995), the use of discourse markers (Sankoff et al. 1997), and have examined community members' reactions to different ways of speaking Montreal French (Thibault & Sankoff 1997). The goals of all of these studies are the same: to describe the competence in French of the first generation of Montreal Anglophones (native English speakers) that had access to French immersion schooling and to find correlations between their linguistic patterns and the following social factors: the first language of the majority of their social network, the language(s) used at work, the amount and type of French they have been exposed to in school and in their community, their scholastic level of French, and sex of the speakers. While we do not directly examine language attitudes, several comments in the interviews indicate that there are attitudinal reasons that the French of Montreal Anglophones differs from L1 French. For example, Lisa, a fluent French L2 speaker who attended a French high school, proudly describes the differences between herself and her Francophone classmates in high school, as illustrated by the quotes in (4-5):

(4) I did not fit in. My accent was different. (Tape 33B)
(5) It's just obvious in my speech [that I'm not French-Canadian]. (Tape 33B)

This is representative of many comments indicating that L2 speakers know that the French they have learned is not, in most cases, identical to L1 Montreal French. This anecdotal evidence is supported by Thibault & Sankoff's (1997) findings from a subjective reaction test, in which a cassette containing excerpts from interviews in French with a number of different speakers (L1 and L2 French speakers) was played to several different audiences. There was a fairly low rate of agreement between the L1 and L2 French auditors: they made similar judgments for only six of 15 speaker/attribute combinations.

---

[2] This paragraph characterizing the speakers is adapted from Nagy et al. (1996), a study analyzing the same set of speakers.

2 Double Subject Marking

We analyze the morpho-syntactic patterns of double marking to see how it corresponds to different manners of language acquisition. This is a process where a subject noun phrase (or pronoun phrase) is "echoed" by an atonic pronoun (Auger 1995). This pattern distinguishes Vernacular Montreal French from the European standard. In standard European French, both patterns exist, but doubling without emphatic intent is much rarer than in Montreal French (J. Auger, p.c.). Therefore, it is not expected that Anglophones who learned their French (only) in school will exhibit it. On the other hand, Anglophones who spend a great deal of time speaking French with Montreal Francophones will have picked up this pattern. Thus, it serves well to separate the effects of school-learning and "street-learning." (6-7) provide examples of two double marking patterns from a corpus of L1 Montreal French (Sankoff-Cedergren 1971 corpus): left-dislocation for emphasis (6) and doubling without pragmatic intent (7). The doubled pronoun is *not* indicative of emphasis or contrast in (7).

(6) Left dislocation (for contrast): **Les maringouins, ils** mesuivent.
Mosquitoes, they *follow me.*
(7) Double marking: **Les maringouins ils** me suivent.
Mosquitoes *follow me.*

(8-10) illustrates that this same pattern exists in the L2 French of Anglophones. These examples are taken from Mike (Tape 37), but all speakers in our corpus exhibit the pattern.

(8) **Moi j**'ai plus de problèmes quand je dis quelque chose puis là je cherche les mots.
Me I *have more problems when I say something, then I look for the words.*

(9) Comme **les cours de français c**'était différentes que le français que mes amis a parlé.
*Like **the French class it** was different from the French that my friends spoke.*

(10) **Mon cours de français c**'était le plus facile.
My French class it was the easiest.

3 Methodology

3.1 Participants in the Study

Our data is taken from interviews with 20 young adult Anglophone natives of the metropolitan Montreal region. The speakers are from two groups, one a self-selecting group of volunteers and the other a sample of the 1990 graduating class of a Montreal area high school. This high school is situated on the border between an Anglophone and a bilingual neighborhood which has received an influx of Francophone speakers in the past several decades. It is an English language high school with two levels of French immersion as well as French taught as a subject (for those less advanced in French). The other participants were recruited by newspaper ads seeking bilingual speakers placed in *Voir* and *The Mirror*, two free Montreal publications. The methodology for interviewing the two types of participants is the same.

All speakers currently live in the Greater Montreal area and speak English with their parents. They differ, however, in their mode of acquisition of French, the type of exposure they had to French as children, their current degree of contact with Francophones, socially and in the workplace, and in the degree to which they use French in their daily lives.

3.2 Interview Protocol

Each participant was interviewed in French by a native Francophone (the second author, M. Fonolossa, L. Gagnon, or G. Sankoff) and later in English by a non-Montrealer Anglophone (the first author). Each interview lasted about one hour. Topics of both interviews include scholastic and family background, use of French in the workplace, attitudes toward French politics, people, and culture, and incidents where language differences have played a significant role. All interviews were tape-recorded and transcribed, providing two types of data: information about the participants' background and a recording of how they speak in both French and English.

3.3 Compiling the Data Sample

Ten English interviews were examined. For each, the first fifteen minutes of the second side of the first tape was examined. All instances of doubled subjects were counted, including emphatics, left-dislocations, etc. The total number of clauses produced in that 15-minute period was counted, and the percentage of double-subjects calculated. Certain types of clauses, which do not permit double marking (e.g., inverted questions, relative clauses, and frozen phrases like "I don't know," "I'm sorry") were excluded.

A slightly different sampling method was used for the French data, since transcriptions were available. Beginning on page 5 of the interview transcription, the first 100 clauses were examined. Each clause was coded for the presence or absence of a doubled-subject and for the independent variables discussed in Section 3.4.[3] No distinction was made between left-dislocated subjects (for emphasis or contrast) and doubled subjects. This decision was made for both pragmatic and theoretical reasons. First, the auditory cues that distinguish the two patterns in L1 French may not be reliably present in L2 speech. A brief pause may distinguish Left-Dislocation from doubling in L1 French, but L2 speech has many more pauses which cannot reliably be distinguished. Similarly, L2 speakers do not use the same patterns of liaison as L1 speakers, so that cue cannot be used. Second, the two patterns may simply be different stages along a grammaticalization continuum, so it is not necessary to consider them as different structures.

3.4 Coding the Independent Variables

No further analysis of the English interview data was conducted, as the goal was simply to illustrate the frequency of doubled-subjects in English. For the French interviews, we considered the linguistic factors in Table 1. These factors were chosen to enable comparison to other research on the same variable (Auger 1995, Nadasdi 1995, Sankoff 1981, Givón 1976). However, the comparisons must be examined with the understanding that the variable is not defined in exactly the same manner in each of these studies.[4]

---

[3] Only doubled subjects at the left edge of the sentence were counted. Any existing right-dislocations were coded as single subjects, since they could not possibly be instances of the double marking phenomenon we are examining.

The dependent variable	
1) Form of the subject	Double marking (dislocation or doubling) Single subject
Independent linguistic variables	
2) Grammatical person	1st - 6th person, masc. 3 sg. Fem 3 pl. fem. 3rd sg. pronoun "ce"
3) Type of subject	Noun Pronoun
4) Clause	main clause subordinate clause
5) Definiteness of subject	Indefinite Definite Non-material
6) Parenthetical between subject and verb	Parenthetic No parenthetic
7) Adverb between subject and verb	Adverb No adverb
8) (Non-subject) clitic between subject and Verb	Clitic No clitic
9) Negative particle between subject and verb	Negation No negation marked there
10) Animacy of the subject	Animate Material but inanimate

Table 1: Linguistic factors

---

[4] Nadasdi (1995) excluded instances of the 3rd person pronoun *ça* in his analysis of double marked subjects in L1 French because of its peculiar behavior. He also excluded all instances of left-dislocation and all ambiguous cases.

Because the envelope of variation differs in each study, it is not appropriate to compare actual values (weights or percentages) across studies. However, the direction of the trends should be comparable.

The social factors considered are the same as those used in Nagy et al. (1996), allowing later comparison of their effects on a phonological and a morphological variable. The list in (11) is, therefore, adapted wholesale from Nagy et al. (1996).

(11) Method of coding "immersion" variables:

Environmental immersion scale

2 points for Francophone spouse, significant other or current roommate
1 point for French friends
2 points for using French at work
1 point if French is the dominant language in a bilingual workplace
.5 point if English is the dominant language in a bilingual workplace

Formal acquisition scale

3 points each for French elementary school or high school
2 points each for an immersion elementary or high school
1 points each for English schools with only the regular French program
1 point for post-secondary education in French

Language used at work

Where both languages are used, the one reportedly used more appears first.

Language of friends

"S" in this column means a spouse, partner or roommate is Francophone.
"F" indicates the speaker has Francophone friends.

French grammatical gender marking score

Subjects were rated on their production of correct gender marking on 20 nouns in an extract of conversation from the French interview.

	Age	Envir. Imm. Scale	Formal Acq. Scale	Lang. at Work	Lang. of Friends	Gender Marking Score	% Double subjects
Women							
Liz	23	6	7	F/E	F	100	5
Sandra	24	6	6	F/E		100	21
Jeanne	22	6	6	E/F	F	95	14
Lynne	24	4	5	E/F	F	95	5
Alicia	21	2	4	F/E	F	80	8
Kathy	25	2	4	F/E		100	5
Janet	21	1	4	E/F		90	4
Tammy	24	1	4	F	S	75	3
Glenda	22	0	4	F/E		85	8
Joan	30	1	2	E/F	S	85	19
Men							
Vincent	26	7	7	E/F	S	100	17
Ted	23	6	7	F/E	S	95	22
Kurt	22	1	5	F/E		80	10
Jack	33	0	4	E		65	4
Mike	23	2	3	E/F	F	75	13
Greg	24	0	3	E/F	F	75	8
Larry	26	0	3	E/F		75	3
Ross	22	0	3	F/E	S	90	4
Peter	20	0	2	E/F		65	9
Don	34	0	2	E		75	4

Table 2: Social characteristics and rates of double marking

3.5 Quantitative Analysis

For the French data, 100 tokens for each of 20 speakers were coded for all the variables mentioned in the previous section. Goldvarb 2.0 for Macintosh, a statistical package making use of a logistic regression algorithm to determine the relative effects of each factor, was employed. Factor weights presented were determined using the "one-level" analysis, but the statistical significance of each group was determined using the "step-up, step-down" analysis. The results are presented in Section 5. For the English language data, no analysis beyond a count of the raw data was conducted, as will be discussed in Section 4.

4 Double Marking in English

Fingers are often pointed at possible interference from L1 when L2 patterns are examined. Therefore, we make a brief digression to examine double marking in English before looking at the L2 French patterns. Double marking without emphatic or contrastive intent is rare in English, but it does exist, particularly in southern American dialects (Wolfram & Christian 1976, Southard & Mullar 1998). We have no independent evidence of its existence in Montreal English, however, (12) provides examples of doubled subjects in English from several of the speakers whose French we are examining. The context indicates that no contrastive or emphatic interpretation is appropriate.

(12) **My friend Martine she**'s French (Greg, Tape 2)
The Québecois they know how to party. (Ted, Tape 61)
My sister she's a music teacher in Joliet. (Vincent, Tape 69)

While most of the speakers produced at least one doubled-subject sentence in English, none produced very many. The range of percentages is shown in Fig. 1. The percentage of doubled-subjects in French is also shown. Little correlation exists between the percentage of double marking in English and in French (Correlation coefficient = 0.47; p>.05). Given the extremely low frequency of doubling in English and the lack of correlation between percentages in the two languages, we discount the English pattern as a source of

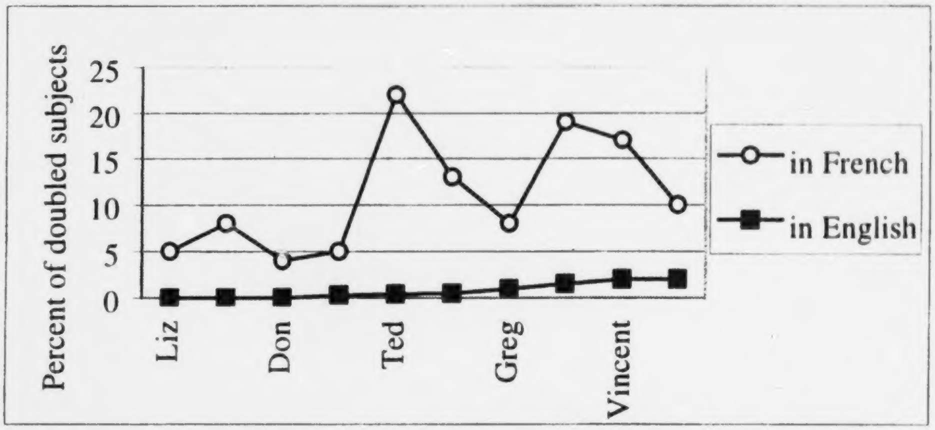


Figure 1: Lack of correlation between double marking in L1 and L2

BEST COPY AVAILABLE

influence on the French pattern and do not consider it as a contributing factor to our model of the French L2 grammar.[5]

5 Results and Discussion

5.1 The Community Grammar for L2 French

No large-scale quantitative study of this variable for native French speakers is available for comparison. What is available is a small-scale study of Montreal speakers (Sankoff 1981) which does not provide quantitative information on the effects of the linguistic variables examined, a study of four Montreal Francophones with some quantitative information (Auger 1995), a descriptive piece by Givón (1976), and a large-scale quantitative analysis examining the speech of Franco-Ontarians (Nadasdi 1995). We examined the effects of as many of the factors studied in these works as possible so that comparison is possible between L1 and L2 speakers.

Our first observation is that the grammar of this group of L2 speakers as a whole resembles the L1 grammar in terms of the effects of a number of linguistic variables. (We can't compare the overall rate of doubling between L1 and L2 speakers because different types of sentences were excluded from the envelope of variation in the different studies.) Table 3 presents the factor weights and frequencies for each variant of the linguistic variables. The application value is ***No** doubling: larger values mean fewer doubled subjects.* The input value is 0.959, indicating a very low overall rate of doubling. In the rightmost column, a "yes" indicates that the effect of the linguistic variable in this L2 sample resembles that reported for L1 data in terms of the direction of the effect (if not the absolute size of the effect).

The first variable examined is grammatical person: number, gender, and referentiality. Overall, we find that there is a higher rate of doubling for third person forms than first and second person, a finding reported by Givón (1976) for L1 French data. (The dotted line divides the high weights from the low weights, and also, more or less, splits the third person singular forms from the others. While third plural feminine falls below the line, its weight is based on only 12 tokens.)

---

[5] Thanks to Yves Roberge for pointing us toward this finding.

Linguistic factor group	**Factor**	**Weight**	**# single /total**	**= L1 effect?**
Grammatical person	"il y a"	0.745	107/108	yes
	1st sg.	0.669	703/719	
	"on" 'nous'	0.681	109/113	
	3rd pl. m.	0.670	130/145	
	1st pl.	0.618	12/13	
	3rd sg. M.	0.332	168/217	
	3rd sg. f.	0.276	108/142	
	"ça"	0.255	420/477	
	3rd pl. f.	0.206	8/12	
Noun or pronoun?	Pronoun	0.562	1582/1675	yes
	Noun	0.176	183/271	
Clause	Subord.	0.787	273/281	yes
	Main	0.445	1492/1665	
Definiteness	Indefinite	0.647	575/618	yes
	Definite	0.430	1190/1328	
Parenthetic (pre-verbal)	No	0.511	1750/1909	yes
	Yes	0.097	15/37	
Adverb (pre-verbal)	No	0.505	1757/1930	yes
	Yes	0.068	8/16	
Clitic *non-sig.* (pre-verbal)	Yes	0.600	128/135	(yes)
	No	0.492	1637/1811	
Negation *non-sig.* (pre-verbal)	"ne"	0.785	41/42	No L1 data
	Ø	0.493	1724/1904	
Animacy *non-sig.*	Animate	0.520	1192/1285	(no)
	Immaterial	0.463	537/612	
	Material	0.436	36/49	

Table 3: Goldvarb analysis to show effects of linguistic factors

The second factor examined is whether the first form of the subject is a noun or pronoun. Native speakers recorded in 1971 produce a higher rate of doubling when the subject is a noun (45%) than when it is a pronoun (10%) (Sankoff 1981). We find the same effect for the L2 speakers: 32% doubling for nouns and 6% for pronouns.

The third factor examined is the position of the subject: whether it is in a main or subordinate clause. Nadasdi (1995) and Sankoff (1981) found fewer doubled forms in subordinate clauses, as did we. The fourth factor, the definiteness of the subject, also produced an effect in our L2 sample that resembled the L1 pattern reported in Sankoff (1981).

The next three factor groups all examine the effect of material intervening between the subject and the verb. In the first two cases, doubling is more likely when material does intervene, both in our data and for L1 speakers (Nadasdi 1995, Sankoff 1981). There is not a significant effect for the presence of other clitics in the L2 data: these L2 speakers do not use preverbal clitics nearly as frequently as native speakers do.

In all cases where all variants of the linguistic variable are robustly present in the speech sample, we find that the L2 speakers have acquired the appropriate effect–the linguistic constraints apparent in the L1 language also exist in the L2 language, illustrating that these speakers have acquired more than just a set of frozen forms that make them sound like Montrealers. The few cases where the effects are not apparent (the last three factors in Table 3) are instances where our sample does not provide enough data to satisfactorily test the effect of the variable. Importantly, there are *no* instances where the L2 speakers exhibit significant contradictory effects of the variables in comparison to the L1 speakers.

5.2 Correlation with Exposure to French

We turn now to the effects of each social variable. Across the different divisions of the group of speakers, the factor weights are ordered as predicted–more contact with L1 speakers leads to more use of the vernacular. The weights and frequencies are shown in Table 4. These values are calculated from the same Goldvarb run as the weights for the linguistic variables presented in the previous section.

We find men use more of the vernacular or non-standard doubled form. We also find that people who use French at work, and therefore have contact with Francophones, have a higher rate of doubling than those who use English at work. Interestingly, the highest rate of doubling is found for speakers who use both languages at work. Perhaps they feel the most pressure to fit in by sounding "local."

There is an interesting effect of the two scales, which, as noted in Sec. 1.4, are strongly correlated. Speakers with higher scores on the Formal Acquisition Scale have higher weights, indicating less doubling, suggesting that the

Group	Factor	Weight	Frequency	# single /total
Sex	female	0.600	0.910	893/984
	male	0.398	0.910	874/964
Work language	English	0.737	0.960	190/198
	French	0.619	0.970	97/100
	both	0.462	0.900	1480/1650
Formal acquisition	5-7	0.558	0.870	583/673
	3-4	0.495	0.940	919/978
	2	0.386	0.890	265/297
Environment scale	0	0.695	0.940	642/681
	1-2	0.453	0.910	627/688
	3-4	0.449	0.840	92/97
	5-7	0.299	0.950	406/482
Grammar grade *[N.S.]*	75-80%	0.532	0.930	633/680
	95-100%	0.507	0.870	593/679
	85-90%	0.459	0.910	356/391
	65-70%	0.444	0.930	185/198
Friendship *[N.S.]*	French friends	0.549	0.890	598/673
	no French friends	0.487	0.920	901/981
	French S.O.	0.434	0.910	268/294

Table 4 : Goldvarb analysis to show effects of social factors

more one studies French in school, the less prevalent are the vernacular features. In contrast, speakers with high scores on the Environment Scale, indicating exposure to French in many non-school environments, have lower weights, indicating more doubling present in their speech.

Not surprisingly, there is no correlation between a person's ability to mark gender correctly on nouns and their frequency of use of doubled subjects, as shown by the fact that this social variable is determined to be non-significant. The scores on the gender test measure a person's scholastic ability–their ability to learn/memorize aspects of the language that *are* explicitly taught in school, in contrast to this untaught variable.

One surprising finding is the non-significance of the effect of having French friends. We would have predicted more use of the doubled forms by speakers with French friends, and especially for speakers with a Francophone partner or room-mate.[6] Overall, however, this set of internal and external factors produces a fairly accurate and comprehensible model of the observed variation. The total χ^2 is 798.

---

[6] The lack of effect may be due to an interaction between this factor group and the Environment Scale group – they represent strongly related factors. However, by that

5.3 Evidence of Acquisition of French Grammar

The final point that we make is that speakers with more immersion in French culture show stronger effects of the linguistic factors that influence the L1. This was determined by looking at the *size of the effect* of each linguistic variable for groups of speakers at different points along the Environment Scale. Table 5 lists the linguistic variables, one in each row. The four middle columns show the average *difference* in the percent of doubling between the two contexts specified by the linguistic variable (e.g., adverb present vs. adverb absent) for all the speakers with scores in each range of the Environment Scale.

Support for our hypothesis comes in the form of larger numbers (bigger differences) in the rightmost columns than in the leftmost columns, indicating that there is a stronger effect of the linguistic variables for the more immersed L2 speakers than the less immersed ones. The rightmost cell of each row indicates whether the data for the corresponding linguistic variable supports the hypothesis that more exposure to vernacular French produces stronger effects of the linguistic variables. Where the differences shown are small, speakers appear not to have internalized the L1 rule regarding the relative frequency of doubling in those contexts. These speakers may be mimicking a surface doubled form without having learned when it is grammatically appropriate to do so.

We find strong support from three of the linguistic variables. In each of the first three rows, the numbers increase from left to right, showing stronger effects of these internal variables for speakers who have more contact with Montreal Vernacular French. In the fourth and fifth rows, there is weak support: for clitics: the numbers are in increasing order, but are not different enough in value to allow us to claim a significant effect; for parentheticals, the numbers generally increase from left to right, but with an inversion of the first two columns, which have very similar values. The effects of the linguistic variables in the last three rows do not support our hypothesis as the values

---

argument, the Work Language factor group should also be non-significant, but it is not.

do not increase from left to right. However, no significant effect was found for Animacy, so the fact that the numbers in this row are not in increasing order cannot be taken as contradicting our hypothesis. (Levels of statistical significance are not provided because the numbers being compared are percentages, which do not lend themselves to statistical analysis.)

Linguistic factor group	**Diff. in % between 2 variants**				**Support for hypothesis?**
	Environment Scale score				
	0	*1-2*	*3-4*	*5-7*	
Adverb	20	24	--	85	Strong
Definiteness	0	0	7	14	Strong
Noun v. pronoun	15	30	39	33	Strong
Clitic	1	4	6	10	Weak
Parenthetical	49	42	97[7]	53	Weak
Main vs. subord. clause	7	9	6	4	No
Grammatical person	8	16	3	16	No
Animacy	5	41	96	17	No

Table 5: Differences in percent of doubling between contexts

Further work is necessary to refine this method. In particular, we would like to be able to compare the effects of the factor weights rather than the frequencies. However, we do not currently have enough data to conduct a Goldvarb analysis on sub-samples. Additionally, while it would be more convincing if we could directly compare the size of the effects to the effects found in L1 speech, that is not possible as no comparable analysis of L1 Montreal speech exists. In spite of these limitations, the method appears promising, and has allowed us to formulate a tentative answer to our third question: Yes, there is an approach to L1 norms as speakers become more immersed in French. More generally, we conclude that this group of L2 French speakers is well on its way to acquiring this pattern of Montreal Vernacular French, and that this progress can be attributed to contact with native French speakers.

---

[7] The large value shown for the speakers with a score of 4 is misleading: there are only 5 doubly-marked tokens with parentheticals for this group.

References

Auger, Julie. 1995. A morphological analysis of Québec Colloquial French pronominal clitics. In A. Dainora *et al.* (eds.). *CLS 31-II: Papers from the Parasession on Clitics.* Chicago: Chicago Linguistic Society.

Givón, Talmy. 1976. Topic, pronoun, and agreement. In C. Li (ed.). Subject and topic, New York: Academic Press. 148-186.

Nadasdi, Terry. 1995. *Variation morphosyntaxique et langue minoritaire: le cas du français ontarien.* Ph.D. dissertation, University of Toronto.

Nagy, Naomi, Christine Moisset, & Gillian Sankoff. 1996. On the acquisition of variable phonology in L2. *PWPL* 3.1:111-126.

Preston, Dennis. 1996. Variationist perspectives on second language acquisition. In Robert Bayley and Dennis Preston (eds.) *Second language acquisition and linguistic variation. (Studies in Bilingualism, Vol 10).* Philadelphia: John Benjamins. 1-45.

Sankoff, Gillian. 1981. Usage linguistique et grammaticalisation: les clitiques sujets en français. In N. Dittmar & B. Schlieben-Lange (eds.) *La sociolinguistique dans les pays de langue romane.* Tübingen: Gunter Narr Verlag. 81-85.

Sankoff, Gillian , Pierrette Thibault, Naomi Nagy, Hélène Blondeau, Marie-Odile Fonollosa, & Lucie Gagnon. 1997. Variation and the use of discourse markers in a language contact situation. *Language Variation and Change* 9.2:191-218.

Southard, Bruce & Al Muller. 1998. Blame it on Twain: Reading American Dialects in *The Adventures of Huckleberry Finn.* In Dalin Oaks (ed.) *Linguistics at Work: A Reader of Applications.* Harcourt Brace. 565-73.

Thibault, Pierrette & Gillian Sankoff. 1997. The insertion of a second language into the community repertoire: Anglophone French in Montreal. Paper presented at *NWAVE 26*, Université Laval.

Wolfram, Walt & Donna Christian. 1976. *Appalachian Speech.* Arlington, VA: Center for Applied Linguistics.

Naomi Nagy
English Department
Hamilton Smith Hall
University of New Hampshire
Durham, NH 03824 USA
ngn@hopper.unh.edu

Hélène Blondeau
Département d'anthropologie
Université de Montréal
C.P. 6128 Succ. A
Montréal, Quebec H3J 3J7 CANADA
Helene_Blondeau@inrs-culture.uquebec.ca

Testing the Creole Continuum

Peter L. Patrick

This paper follows earlier work presenting a variationist model of the creole mesolect (Patrick 1996) and enriching the speech community concept so that it was capable of handling Caribbean sociolinguistic complexes (Patrick 1998a). Here I re-evaluate the creole continuum's usefulness in accounting for variation in an urban creole, focusing on the criterion of discreteness. Forthcoming work also evaluates the continuum concept, considering the criterion of unidimensionality, i.e. the correlation of linguistic variation with social factors.

The series of papers largely review the same data, which I gathered — both in sociolinguistic interviews, using Jamaican Patwa, and also in more formal interviews and test environments, using Standard Jamaican English — from residents of an urban neighborhood of Kingston, Jamaica (pseudonym 'Veeton') in 1989-90. These data are only briefly summarized in the present paper, which sketches patterns across a range of individual speakers; they are more fully analyzed in Patrick (1998b). Below I take for granted a number of points argued for there and elsewhere.

The classic variationist notion of a speech community requires:

- that a social community exist,
- that its members share patterns of language use, and
- crucially, that they share norms for social evaluation of speech.

In studies of New York City (Labov 1966), Detroit (Wolfram 1969), Belfast (Milroy 1980), Sydney (Horvath 1985) and many other urban areas, we find that individual variation is distributed differentially in a manner that is partially constrained by linguistic factors, is partially ordered by social stratification, and is interpretable in the light of values which are to a great degree held in common. Guy (1980) showed convincingly that the speech of individuals in the New York and Philadelphia speech communities reflect group norms and confirm this model. The creole continuum in its original formulation (DeCamp 1971, Bickerton 1973) was held to challenge this notion of the speech community. In fact, Guy's 1980 article was a direct response to Bickerton's, which challenged the whole enterprise of grouping individual speakers together. Bickerton charged that such aggregates were in general an illusion and that the individual norms they represented would, if investigated, prove to be an unruly lot with no real coherence; Guy, looking at (TD)-

deletion, showed that this need not be true.

Grammatical rules in a continuum situation, it was said, differ significantly between individual speakers in such a way that they cannot be resolved into discrete social or geographical dialect groupings, because the differences refuse to bundle together and require a large variety of norms to account for all the individual behaviors. At the same time, the contrast between the extreme ends of the continuum was taken to be too great to allow for shared patterns of use or, presumably, for a unified evaluative mechanism. In the 'polylectal' model offered by Bickerton and C.-J. Bailey (1973), there *is* a linear ordering of varieties — but it is along a purely linguistic dimension, not predicated upon the social stratification of speakers. In fact, DeCamp (1971) and Bickerton (1980) stress that social characteristics of speakers should not be used *a priori* to distinguish varieties (though, at least in DeCamp's work, it is understood that the linguistic ordering of lects can be correlated with (one or more) social dimensions).

The two larger problems here — the grammatical structure of mesolectal speech and the sociolinguistic structure of the urban creole speech community — can be viewed independently for analytical purposes, though I believe they are ultimately inseparable. The first three points have been recast by Rickford (1987) as the criterion of (non-) discreteness. Basically, the claim is that variation across speakers is very fine-grained and covers a lot of ground, so that the traditional solution of consolidating them into distinct varieties, or dialects united under a single umbrella, is unworkable — in part because of the structural distance between the most standard (or acrolectal) and deepest creole (or basilectal) grammars. This is the topic of the present paper.

The fourth and fifth criteria are rephrased by Rickford as the requirement of unidimensionality: the variation found across the continuum can be expressed along a single dimension. Elsewhere I have argued (Patrick 1998b) that such a dimension cannot be purely linguistic in nature, but must be social or sociolinguistic; and that the number of dimensions required to model variation must be an empirical and historically contingent matter, not a theoretical requirement. This leaves only discreteness or continuity as the essential characteristic of a creole continuum (though if it could not be correlated with, and explained by, one or more social dimensions, the construct would be of little use). I will not pursue the correlation issue here.

The question here is whether the creole continuum can be maintained as a descriptively useful concept, and with what modifications. Considering the urban Jamaican speech of Veeton, what conclusions about discreteness can we draw? Are there distinct varieties? a true continuum? or what Bailey (1974) has called a 'gradatum' — a situation where there may be continuous variation *within* a wide mesolect, but sharp boundaries *between* it and the acrolect and basilect?[1] Or is some other solution required?

This immediate descriptive problem aside, I will argue that the gap which earlier researchers believed to exist between complex creole communities and variationist methods of analysis can and must in fact be bridged. The urban JC mesolect consists of a mixed, variable grammar featuring the systematic presence and integration of English forms and rules, distinguished by their level and patterns of variability (Patrick 1996). It is also independently necessary (Patrick 1998a) that the concept of the speech community be expanded to allow for the multiple, mutually-opposing sets of norms that characteristically co-exist for (synchronically-) related varieties in a creole continuum (as in other complex communities, Santa Ana and Parodi 1998).

This paper continues the project of reconciling creole speech data and practices with the principles of language variation and change by updating and modifying the notion of the continuum. This is in answer both to the relative neglect of the challenge that creole data pose to variationists (or their assumption that such data must constitute a fundamentally different sort of sociolinguistic structure, Guy 1980), as well as in answer to some creolists' rejection of both the continuum and/or variationist analysis as incapable of modelling the richness and heterogeneity of Caribbean sociolinguistic structures (e.g. Carrington 1993, Mufwene 1994, Winford 1988, 1993).

The continuum was first proposed by David DeCamp (1971) to give a unified account of the Jamaican sociolinguistic situation, which he found to be poorly modelled by structural dialectology, diglossia, and other dichotomous descriptive methods of the time. Any empirical solution must be based on some selection of particular linguistic phenomena; and DeCamp's answer, affirming the continuum, was exemplified by his analysis of a variety of unrelated lexical items and surface phonological variables.

But the case for the independence of Jamaican Creole from English rests on the analysis of deeper grammatical properties involving central elements of the verb phrase, such as Tense-Mood-Aspect marking systems. DeCamp's data illustrated his belief in continuity, but it did *not* constitute a test of the continuum model — he did not conduct a sustained examination of appropriate features.

Accordingly, the present analysis includes two types of past-reference marking — a creole-like one with preverbal '*did*' and '*neva*', and an English-like one using verb inflection — as well as a related phonological variable,

---

[1] This distinction is implicit in Bailey's 1974 discussion; Fasold (1990:196) raises it explicitly.

(TD)-deletion — simplification of final consonant clusters ending in an alveolar stop — which is motivated quite independently of past-marking but intersects with it. The other phonological element analyzed, (KYA) or palatalization of velar initials before low vowels, has no connection with the syntax, and is thus similar to the surface-level variation of DeCamp's items. The variation profile that emerges will tell us more about the grammar than DeCamp's study did — if non-discreteness indeed characterizes it, it will not simply be because we are looking at unrelated elements which have no reason to cluster together.

Figures 1 and 2 show the behavior of individuals in both interview and formal-test situations. All speakers recorded several hours of sociolinguistic interviews; later, they also performed English-to-Creole and Creole-to-English translation tasks, in which I presented them with tape-recorded stimuli — five or six sentences, one at a time, which formed a brief narrative set in the past — and asked them to translate each sentence into the other variety (if it was English, translate into Patwa; if it was Patwa, then into English). Each task was loaded with opportunities for speakers to utter variant forms of the linguistic variables of interest, namely (TD)-deletion and (Past)-marking.

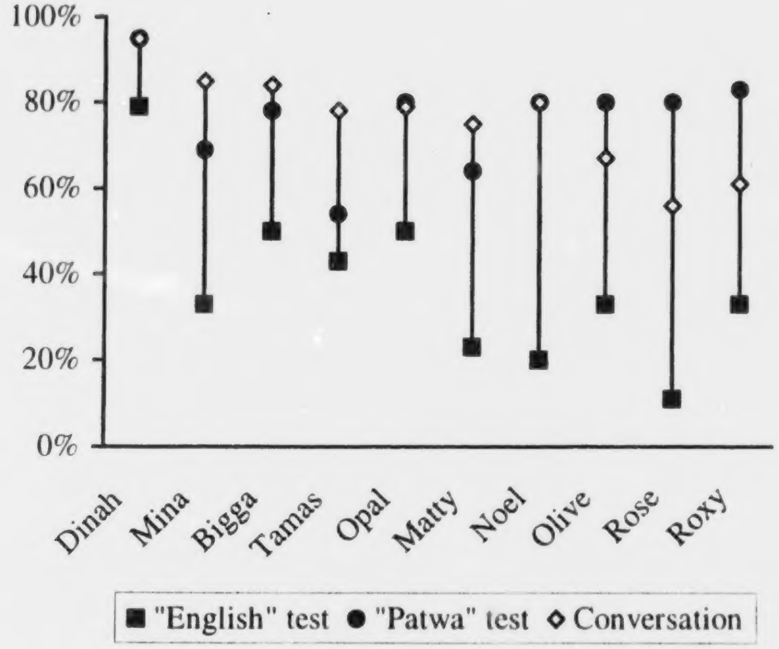


Figure 1: Test and conversational data for (TD)-deletion

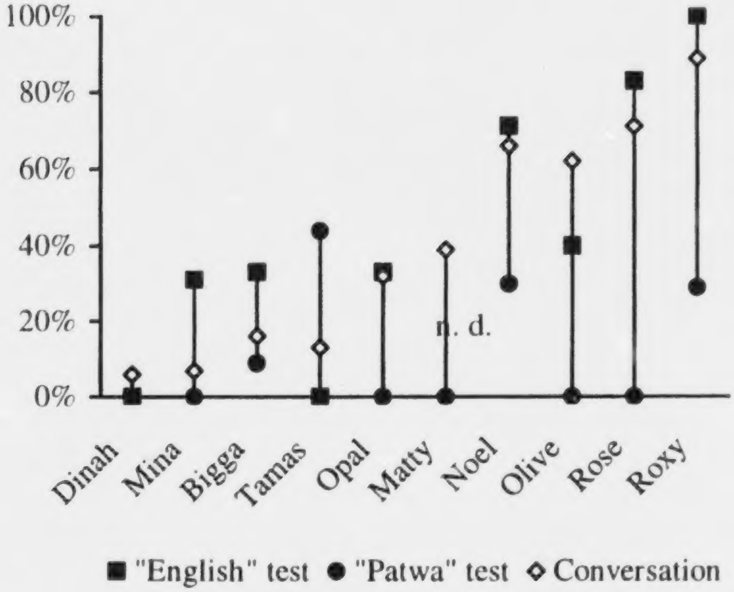


Figure 2: Test and interview data for combined (Past)-marking

In both Figures 1 and 2, an empty diamond shows a speaker's performance over the length of interview data recorded with them. (These interviews took place on separate occasions, generally long before the tests.) A filled square shows how speakers behaved when asked to translate Patwa sentences into English; a filled circle shows translations from English into Patwa. These tasks, then, allow us to contrast speakers' production in a relatively natural situation with their stereotypes of Patwa and English targets.[2]

I have earlier (Patrick 1998a) used these data to illustrate the simultaneous existence of opposed evaluative norms — some unified and some dichotomous — within the same speech community. Here the issue is whether individual variation is or is not continuous: do speakers fall into subgroups exhibiting fine or sharp stratification? Where quantitative differences in interview speech appear, do they coincide with contrasting targets/stereotypes?

The interview data for (TD) — the diamonds in Fig. 1 — show continuous variation with fine stratification. There are no sharp breaks, and there is a regular decline in (TD)-absence from left to right as status rises. There appears to be a single norm, in the sense that in interview speech the level of

---

[2] Social characteristics of speakers appear in Table 3. Note that speakers in Figs. 1-2 are stratified by status, low at the left and high at the right. Note also that not all of the 16 speakers listed in Table 3 appear for all variables (see Patrick 1998b).

BEST COPY AVAILABLE

this variable runs closer to the "Patwa" target than the "English" one; while this is attenuated for the highest-status speakers, Roxy and Rose, the difference is gradual and not dramatic. In general, cluster-simplification is very high — always over 50%, and at or above 75% for most speakers — and for the lowest-status speaker, Dinah, it approaches 100%.

The interview data for (Past)-marking in Fig. 2 also show fairly continuous variation. The range covers most of the spectrum, from 6% to 89%. The sharper gradient suggests that this grammatical variable may be a more salient, if still quantitative, marker of social status. Stratification into three groups (Low, Mid and High) is apparent: the Mid group contains two speakers, Opal and Matty. But what distinguishes the mid- and high-status speakers in the sample from the low group is that their rate of verb inflection in interviews, instead of reflecting their "Patwa" norm, is closely identified instead with their "English" target (consider speakers from Opal over to Roxy).

The third variable (KYA), palatalization of velar initials in words such as /gyaadn/ 'garden' and /kyap/ 'cap', shows a completely discontinuous picture. Figure 3 demonstrates that the variation falls largely along age-group lines. All young speakers (under 25), as well as the oldest speaker (Rose, 82), show categorical absence of the variable in one environment — before historically AR words ('card, guard') — in which the other speakers have it very frequently (68% for the Older and 90% for the Middle-aged, respectively). In other environments, especially Short-A words, the variable occurs for all speakers at essentially categorical rates. This is due to a change in progress towards a pre-existing prestige pattern long held by speakers like Rose, but now being acquired by the urban young of all classes through education and mass media (Patrick 1998b). The two norms are quite dichotomous — there is nothing resembling continuous variation here:

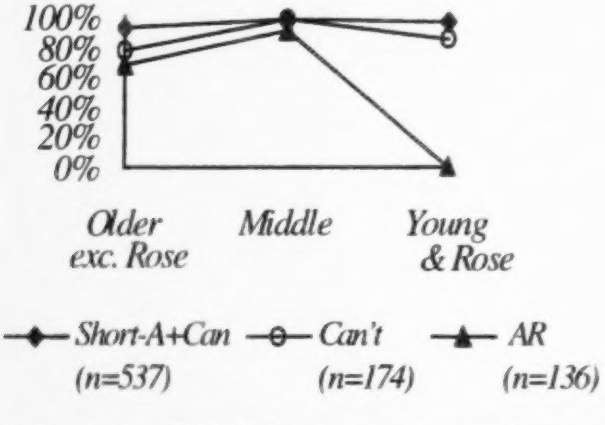


Figure 3: Interview (KYA) data by word-class & age (except Rose)

For pre-verbal past-marking with '*did*' there is also a qualitative distinction: though all speakers recognize and properly interpret the form in translation tasks from Patwa, only older people use it in conversation regularly. Of 100 forms found in some 2,000 past-reference clauses, 96 were used by four speakers over 45; the fifth older speaker used it twice (this was Rose, who showed the prestige pattern for KYA), and two younger speakers used it once each. This can be seen in Table 1 below, in the column labeled '*did*'.

The same sharp divide can be seen with negative past marker '*neva*'. The form is used by all speakers, but only a few — Roxy, Rose, Noel, Olive — use it like the familiar English adverb, independent of the {-ed} tense inflection on the verb. For all other speakers, '*neva*' is a classic creole preverbal tense marker: the following verb never carries the redundant {-ed} inflection.

So it is sometimes possible, on purely linguistic grounds, to consolidate speakers into subgroups that contrast with each other yet show internal agreement on key properties — quantitative norms, as well as underlying principles and patterns of linguistic organization. But what are we to make of the fact that these subgroups differ from variable to variable?

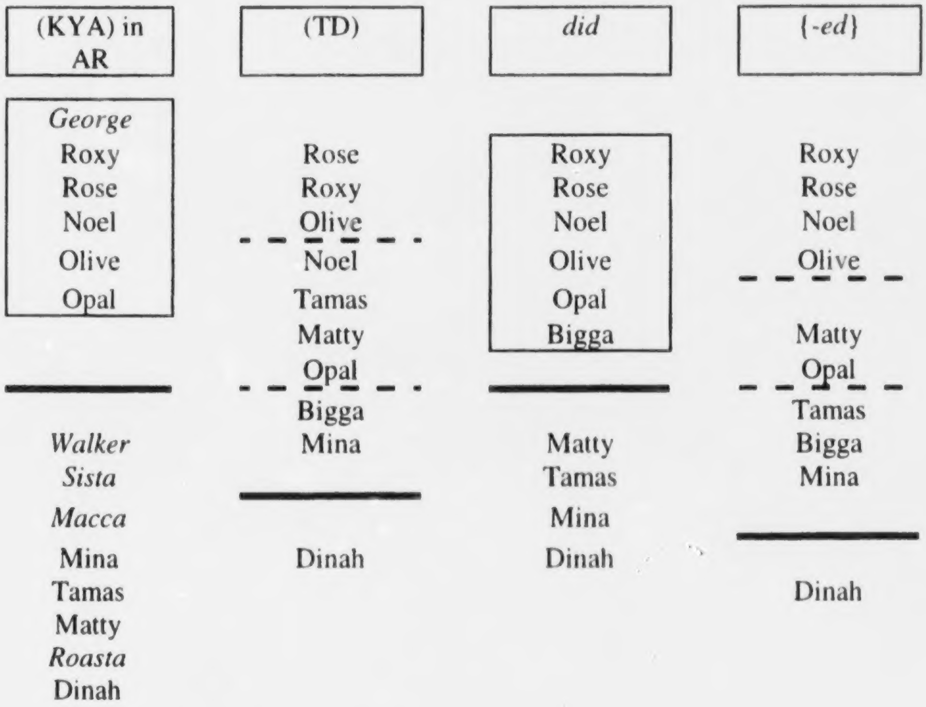
Figure 1 gave no grounds for distinguishing subgroups at all, but rather resembled DeCamp's (1971) illustration of a classic continuum; Figure 2 suggested three subgroups, finely stratified and correlated with status; while Figure 3 clearly requires 2 subgroups, largely distinguished on the basis of age, but falling into groupings incompatible with those suggested by Figure 2. One assumes that with these data we have not yet reached the limits of variability. These results raise several questions. First:

- Does the hierarchy or ranking of speakers fluctuate across subgroups?
- If so, what explains the relative position of speakers?

(These answers pertain to unidimensionality and the sociolinguistic structure of the speech community, rather than the current issue of discreteness — the organization of the grammar.) Secondly,

- Do the subgroups constitute distinct varieties?
- Do grouping patterns persist across variables, i.e. across the grammar?

To answer that, it will be helpful to compare the patterns across all variables at a glance. In Table 1 (next page) I have only represented the AR environment for (KYA), because there is no variation in other ones; and I have split (Past) into its two components, classic pre-verbal creole marker '*did*',



(KYA) in AR	(TD)	*did*	{-*ed*}
George			
Roxy	Rose	Roxy	Roxy
Rose	Roxy	Rose	Rose
Noel	Olive	Noel	Noel
	- - - - -		
Olive	Noel	Olive	Olive
			- - - - -
Opal	Tamas	Opal	
	Matty	Bigga	Matty
	Opal		Opal
———	- - - - -	———	- - - - -
	Bigga		Tamas
Walker	Mina	Matty	Bigga
Sista		Tamas	Mina
	———		
Macca		Mina	
			———
Mina	Dinah	Dinah	
Tamas			Dinah
Matty			
Roasta			
Dinah			

Table 1: Discontinuities in the distribution of 4 linguistic variables

versus English-like inflection with {-*ed*}. (The data for negative past '*neva*' do not appear.)

Speakers who only appear once are given in italics; I represent categorical patterns in conversational data by enclosing the speakers who exhibit them in a box. All speakers not so enclosed are vertically ordered by their probability of use of the variable. Columns are arranged so that usage approximating the standard appears at the top. Dotted lines, such as between Olive and Matty for the {-ed} variable, represent smaller quantitative discontinuities; solid lines represent greater quantitative and/or qualitative ones.

Remember that all these speakers were drawn from the middle levels of Jamaican society, and that their speech is also intermediate — so we want to pay attention to the extremes. Are there clearly distinct speakers: those consistently among the most standard, who should belong to the acrolect, or those consistently most creole-like who should be classed with the basilect?

Rose and Roxy are the most consistently standard, but their performance in conversational speech is never very far above that of their near neighbors. In Figures 1 and 2 it's clear that they share identical target norms for 'Patwa and 'English' with other speakers. It might be just possible to differentiate Roxy from the rest for Past inflection with {-ed}: she has above 80% inflection in Past clauses, produces the form categorically in the 'English' test situation, and inflects Irregular verbs at 96% (not shown here). But this is the only variable for which she might be singled out; on the rest she seems simply to be at the top end of this mesolectal sample.

At the bottom end, the situation is quite different. Compare Dinah and Mina, the most non-standard speakers in Figs. 1 and 2. Dinah shows nearly complete absence of (TD) in conversation and utterly misses the 'English' target — the other speakers are variable here too, but she shows 80% absence while they are all under 50%. For {-ed} things are even clearer: these two speakers have similar, and very low, rates of inflection in conversation, but like everyone else Mina clearly distinguishes the 'English' and 'Patwa' stereotypes, while only Dinah has no clue — she cannot inflect regular verbs even in the test situation.

Dinah shares the traditional, non-prestigious pattern for (KYA), and the creole use of *'did'* as a pre-verbal TMA marker, with other speakers. For the latter, note that all users of this form (and all who use *'neva'* in a similar way) are subject to the classic constraints of anteriority and punctuality identified by Bickerton (1975) — see Table 2.

	did/neva + Verb		Verb + {-ed} inflection			
	±punc	±ant	±punc	±ant	prec. envt	foll. envt
Roxy	...	...	∅	∅	❶	✸
Rose	...	...	∅	∅	❶	✸
Noel	...	...	∅	∅	❶	✸
Olive	...	...	∅	∅	❶	✸
Opal	...	...	∅	❶	∅	✸
Matty	✸	❶	∅	❶	∅	✸
Bigga	...	...	✸	∅	❶	≈
Tamas	✸	❶	✸	∅	❶	≈
Mina	✸	❶	✸	∅	❶	≈
Dinah	✸	❶	✸	∅	❶	≈

❶ Primary constraint ✸ Secondary constraint ≈ Equivocal constraint
∅ Not significant ... (Not applicable)

Table 2: Relative influence of constraints on past-marking, by speaker

What this asymmetry shows us is that traditional creole elements and constraints do indeed intrude into the mesolect, though they are not used actively by all mesolectal speakers and tend to be restricted to those who are

Speaker	Age	Years of School	Occupation and Social Class
George	23	16	Civil servant, MS-2
Roxy	14	*(9…)*	*(still in school, → MS-2?)*
Rose	82	11 +	Head nurse (ret.) , MS-2
Walker	70	11 +	Headmaster, school (ret.), MS-2
Matty	49	12 +	Photographer, MS-3
Macca	72	6	Police detective (ret.), WC-1
Olive	24	12 +	Accounts clerk, MS-4
Noel	17	11	Clerk-trainee, MS-5
Roasta	30	12 +	Toolmaker, WC-2
Opal	17	11	*(looking for work, → MS-5?)*
Tamas	70	3	Shoemaker, PB-2; also factory worker, WC-2
Mina	75	6	Dressmaker, PB-2
Sista	32	3	Self-employed actor, activist, PB-2
Bigga	17	*(10...)*	*(still in school, → WC-?)*
Dinah	46	3	Domestic helper, WC-5

MS=Middle Strata; PB=Petit Bourgeoisie; WC=Working Class (Gordon 1987)

Table 3: Social characteristics of speakers (by occupational status)

older and/or of lower status. But if the occurrence of *basilectal* elements cannot reliably identify a speaker, the *absence of English* forms, constraints and targets does — and by this criterion Dinah is clearly not a mesolectal speaker, but belongs to the basilect.

Generalizing from our findings about individual variation here, and putting them together with earlier work on the nature of the mesolect and the evaluative mechanisms at work in constituting the creole speech community, I will now go back and answer the earlier question about discreteness and the continuum. Since the variable presence and systematic integration of English forms and rules defines the mesolect, there appears to be no clear dividing line *in the grammar* between the mesolect and the acrolect — a speaker like Roxy may reach the categorical level in one variable while not differing essentially from her peers. But the absence of such knowledge boldly marks off basilectal speakers.

Thus the situation in Jamaica most closely resembles Bailey's notion of *gradatum*: continuous variation within a wide mesolect, but a sharp boundary on the lower end, between it and the basilect. The extent and type of grammatical variation shown above is indeed greater than customary in classic standard and dialect situations, but it does not require an unmanageable number of lects and it does fall into a limited number of shared patterns of use. Taking into account these modifications and others noted earlier, I conclude that the revised notion of the creole continuum is a rich, flexible tool capable of accommodating variationist methods of analysis and modelling the linguistic variation in Jamaican speech.

References

Bailey, Charles-James N. 1974. "Some suggestions for greater consensus in creole terminology." In DeCamp, David and Ian Hancock, eds., *Pidgins and Creoles: Current Trends and Prospects*. Washington DC: Georgetown University Press: 88-91.

Bailey, Charles-James N. 1973. *Variation and linguistic theory*. Arlington, VA: Center for Applied Linguistics.

Bickerton, Derek. 1980. "Decreolization and the creole continuum." In Valdman, Albert and Arnold Highfield, eds., *Theoretical Orientations in Creole Studies*. New York: Academic Press: 109-28.

Bickerton, Derek. 1975. *Dynamics of a creole continuum*. Cambridge: Cambridge University Press.

Bickerton, Derek. 1973. "The nature of a creole continuum." *Language* 49(4):640-69

Carrington, Lawrence D. 1993. "Creole space — a rich sample of competence?" *Journal of Pidgin and Creole Linguistics* 8(2):227-236.

DeCamp, David. 1971. "Towards a generative analysis of a post-creole speech continuum." In Dell Hymes, ed. 1971, *Pidginization and creolization of languages.* Cambridge: Cambridge University Press, 349-70.

Fasold, Ralph. 1990. *The sociolinguistics of language.* Oxford: Blackwell.

Gordon, Derek. 1987. *Class, status and social mobility in Jamaica.* Kingston: Institute for Social and Economic Research.

Guy, Gregory R. 1980. "Variation in the group and the individual: the case of final stop deletion." In William Labov, ed., *Locating language in space and time.* New York: Academic Press, 1-36.

Horvath, Barbara M. 1985. *Variation in Australian English: The sociolects of Sydney.* Cambridge: Cambridge University Press.

Labov, William. 1966. *The social stratification of English in New York City.* Washington, DC: Center for Applied Linguistics.

Milroy, Lesley. 1980. *Language and social networks.* Oxford: Blackwell.

Mufwene, Salikoko. 1994. "On decreolization: The case of Gullah." In Marcyliena Morgan, ed., *Language and the social construction of identity in creole situations.* Los Angeles: Center for Afro-American Studies Publications, UCLA, 63-99.

Patrick, Peter L. 1998a. "Caribbean creoles and the speech community." Paper given to the Society for Caribbean Linguistics XII meeting, Castries, St. Lucia W.I. August 1998.

Patrick, Peter L. 1998. *Urban Jamaican Creole: Variation in the mesolect.* (Varieties of English Around the World, 17.) Amsterdam/Philadelphia: Benjamins.

Patrick, Peter L. 1996. "Insights into the mesolect." Paper given to the NWAVE-XXV meeting, Las Vegas NV, October 1996.

Rickford, John R. 1987. *Dimensions of a Creole Continuum.* Stanford University Press.

Santa Ana, Otto, and Claudia Parodi. 1998. "Modelling the speech community: Configurations and variable types in the Mexican Spanish setting." *Language in Society* 27(1):23-51.

Winford, Donald. 1993. *Predication in Caribbean English Creoles.* (*Creole Language Library*, 10.) Amsterdam, Philadelphia: Benjamins.

Winford, Donald. 1988. "The creole continuum and the notion of the community as locus of language." *International Journal of the Sociology of Language* 71:91-105.

Wolfram, Walt. 1969. *A linguistic description of Detroit Negro speech.* Washington DC: Center for Applied Linguistics.

Department of Language & Linguistics
University of Essex
Wivenhoe Park
Colchester CO4 3SQ
United Kingdom
patrickp@essex.ac.uk

Going Younger To Do Difference:
The Role of Children in Language Change

Julie Roberts

1 Introduction

Some of the critical questions explored in research on language change are those of who initiates the changes, who participates, and who retains more conservative forms. Gender, race, socioeconomic class, and ethnicity have all been widely studied in terms of their effect on linguistic change. While the age of the speaker has also been found to be important in predicting their participation in change in progress, it is only recently that the age range examined has been extended to include preschool children. The current study seeks to explore the findings of these studies of preschool children, as well as some new data, in order to examine the ways in which these young children can participate in language change and the possible learning mechanisms which enable this participation.

1.1 Role of Children in Language Change

Because most of the earliest work on language change centered on adolescents through adults, the first step in extending this range was to ascertain whether or not children participated at all in this phenomenon. To this end, Roberts and Labov (1995) examined the acquisition of the complicated Philadelphia short *a* pattern in 3- and 4-year-old children. They found that the children were well into the process of learning this pattern. They had acquired some of the constraints on raising and tensing short *a*, such as in environments preceding nasal consonants and in the words *mad*, *bad*, and *glad but not* the similar word, *sad*. They were in the process of acquiring other features of the pattern, such as raising before voiceless fricatives and affricates, in that the 4-year-olds were significantly more accurate than 3-year-olds in their raising of short *a* in these environments. Finally, and most critically to the current questions, the children extended the pattern of short *a* raising by producing the tensed and raised vowel significantly more often than their parents in a new, lexically conditioned raising environment – the word *planet*. These findings support the argument that children not only acquire even the most complex parts of their dialect at a very young age, but they also actively participate in changes by extending their parameters.

In an extension of this study, Roberts (1997a) examined the role of language input on children's participation in change in progress. All of the variables in question in that study were undergoing change in progress and varied in the complexity of this change. (See Payne 1990, for a description of Philadelphia vowels undergoing change and their complexity.) All of the children were found to have acquired the simpler, phonetic changes (e.g., fronting of (aw) as in *cow*). On the other hand, more complex changes (e.g., raising of checked (ey) as in *cake* and raising and tensing of short *a*) were acquired only by those children who had parents who were native Philadelphians. Finally, none of the children had acquired the centralization of long (ay) before voiceless consonants, as in *kite*. This change had been found to be led by male speakers in Philadelphia (Labov 1990), so it was not surprising that the children, whose child care situation consisted primarily of women – their mothers and female child care workers – had not learned this change. These findings supported the prediction made by Labov (1990) that in a community in which child care was provided mostly by women, female-led changes would be accelerated, as opposed to male-led changes.

Although these findings do indeed underscore the participation of children in their speech community and the importance of including them in its description, the question remains as to how extreme the individual vowel productions of these children are relative to those of their parents and other adults in their community. Kerswill (1996) provided some insight into this issue in his study of the acquisition of a new dialect in Milton Keynes, Great Britain by preschool children through adolescents. Using a methodology based on researcher judgement coding, he found that although in general the productions of preschool children were less extreme than those of adolescents, there were some important individual variations. Two of his 4-year-old speakers' productions were similar to those of older peers; one child's productions were much like those of his father; and one child produced forms that appeared to reflect a compromise strategy between his parents' productions. These findings underscore the importance both of input in dialect acquisition and that of data interpretation encompassing individual differences as well as group trends.

The advent of normalization techniques allows the direct comparison of computer analyzed vowel tokens which could previously only be compared qualitatively as complete systems (Neary 1977). Therefore, one of the purposes of the current study is to examine the normalized data of the Philadelphia children discussed above to see if their vowel productions match those of their parents.

1.2 Role of Input on Children's Early Phonological Acquisition

The role of language input directed to children by parents and other care takers has been widely researched by psycholinguists. Those favoring a more nativist explanation of child language acquisition minimize the importance of this child-directed speech (CDS), noting that minimal input is sufficient for a child to acquire a full linguistic system at an amazingly fast rate (Pinker 1984, Gleitman and Wanner 1982). Others, noting both the ways in which those producing CDS simplify, shorten and otherwise tailor their language to the child learner, argue that this language is specifically fitted to the language acquisition process and has an effect on the rate of acquisition, at minimum, and perhaps the style and quality of acquisition as well (Murray, Johnson, and Peters 1990, Clarke-Stewart 1973).

As we explore the child's participation in language change, it is important to consider what the possible mechanisms are that allow or promote the acquisition of dialect features. Related work on the acquisition of language-specific phonology is somewhat more limited than that concerning morphology and syntax, but Patricia Kuhl and colleagues have looked at maternal vowel input to infants acquiring English, Russian and Swedish (Andruski and Kuhl 1996, Kuhl, Andruski, Chistovich, Chistovich, Kozhevnikova, Ryskina, Stolyarova, Sundberg, and Lacerda 1997). Using acoustically analyzed natural language input of mothers to 2- to 5-month-old infants and to adult interviewers, they found that the vowels measured, (long (iy), (ah), and long (uw)), were significantly more peripheral (i.e., toward the edges of the vowel space) when addressed to infants than to adults. This resulted in an "expanded vowel space, one that is acoustically 'stretched'" (Kuhl et al. 1997:84). The 'stretching' of the vowel triangle was substantial in all three languages studied – 91% for English, 94% for Russian, and 90% for Swedish. The authors concluded that the vowels were therefore more distinct from each other providing greater featural contrast among vowels while minimizing acoustic overlap. This well-specified vowel information would ultimately aide phonological acquisition. A further study revealed that infants attended to and increasingly imitated these productions (Kuhl and Meltzoff 1996).

Whereas the arguments made by Kuhl and colleagues center on the process of language acquisition regardless of the particular language involved, they raise intriguing possibilities for those of us interested in the learning of specific dialect features as well. For example, it seems reasonable to hypothesize that this more extreme vowel input, which children receive during infancy, could become the targets of their own early productions. Therefore, in the cases in which the innovative dialect variant is also the more peripheral variant, these early CDS productions might become the children's typical productions as they mature, resulting in a more peripheral, or extreme, innovative form. Hence, the language-learning situation could be seen as specifically aiding the children in their participation in some aspects of language change (i.e., those involving peripheral vocalic movement).

Of course, there are many unanswered questions raised by these studies. Among them, does this peripheralization of CDS vowels at the 'points' of the vowel triangle extend to other vowels as well, particularly those involved in change in progress? What are the acoustic features for vowels during CDS in which the changing target is moving away from the periphery (i.e., fronting of Philadelphia (ow) and (uw))? Do the children's fathers and other male care takers also participate in this phenomenon, and, if not, would this further the deceleration of certain changes, as suggested by Labov (1990) and Roberts (1997a)? How long does the vowel triangle stretching take place, and do the children's first vowel productions match these 'stretched' targets? These are questions necessitating several studies, but it is an additional purpose of the current one to begin this process by examining the productions of 3- and 4-year-old children and their mothers and comparing vowel extremity, both in terms of changing Philadelphia changing vowels and the three vowels examined by Kuhl.

1.3 Social Interaction and Children's Role in Language Change

Social interaction has played a prominent role in both developmental accounts of child language acquisition and in variationist accounts of language change, but in both cases, the traditional emphasis has been on the interaction produced by speakers as a reflection of particular immutable biological categories (i.e., sex, race, ethnicity, age). (See Eckert 1990 and Thorne 1993 for discussions of this issue.) Similarly, in child socialization literature, children have been viewed as the receivers of local cultural norms from their care takers by means of particular socializing strategies, many of them linguistic (Schieffelin 1990, Ochs 1988, Heath 1983).

West, Zimmerman, and others have argued that these categories are not immutable but rather socially defined and constructed. Community members can be described not as demonstrating or reflecting race, sex, age, etc. in their behavior but as constructing their identities on a day-to-day basis or "'doing' difference" (West and Zimmerman 1987, West and Fenstermaker 1995). Eckert (1989, 1988) has applied this approach to variationist research and argued that gender is constructed by women through their adoption and use of certain dialect forms as through other forms of behavior. The position taken in the current study is that children, too, 'do' community from their earliest acquisition and use of dialect features and that this behavior is negotiated and constructed beginning in the early care taker-child interactions and continued later through peer interactions. As shown in Roberts and Labov (1995) and Roberts (1997a), these early interactions can result in the influencing of the parameters of change and, possibly, the favoring of some changes over others. In other words, these earliest interactions could initiate the child into the possible domains of language change. The final purpose of the current study will be to see if the comparison of the extremity of vowel tokens suggests that children's early forays into the domains of change may result in their having an effect on the ultimate targets of change.

1.4 The Current Study

In summary, the study of childhood dialect variation is, like its subjects, in its infancy. There are many questions to be answered and issues to be explored. The overall purpose of the current study is to bring together some of the information, theoretical and empirical, from several sources which may bear fruit in this research. It is critical not only that we document children's acquisition of dialect forms and their participation in speech communities, but also that we begin to explore the mechanisms by which this participation takes place. The findings from the fields of adult language variation and change, psycholinguistics, and social interaction appear to hold much promise in the future of this research. The data to be presented here is intended as an initial step in the integration of some of this research and an exploration of the importance it may have in the interpretation of some normalized data on Philadelphia preschool children and their mothers. It is hoped that the resulting analysis will suggest other avenues of study in this area.

2 Methodology

2.1 Participants

Six children, aged 3;4 to 4;11, were tape recorded over a four-month period at their day care center, Kid's Land, located in a working class to lower middle class area of South Philadelphia. The group comprised three girls and three boys, all of whom were native to Philadelphia. Four of the children, Evan, Jenny, Shelly, and Danny, had parents who were also Philadelphia natives. The children with non-native parents were Gia, 3;11, who lived with her father, a Philadelphia native, and her mother, who had moved frequently throughout her childhood, and Mike, 3;4, whose parents were born and raised in Italy and spoke Italian at home.

The primary care takers of the four children with native Philadelphia parents also participated in the study. These included the mothers of Evan, Shelly, and Danny and the paternal grandmother of Jenny. For further information about these participants, please see Roberts (1996, 1997ab).

2.2 Procedure

Play-interview sessions, described in Roberts (1996, 1997ab), were used to gather the child data. The adults were interviewed using sociolinguistic interview techniques (Labov 1984). All sessions were recorded using a Nagra tape recorder and a Sony ECM 50 lavaliere microphone. Single words taken from the spontaneous speech of the participants were digitized using the Kay Computerized Speech Lab (CSL) on a 386 personal computer. Single measurements of the first two formants (F1 and F2) of the vowels were taken as follows. For vowels containing both nuclei and glides (e.g., (iy), (ey), (ow), (aw), (uw), (ay), and tense short *a*), nuclei were measured. Specifically, if the extremity of raising or fronting was of interest, the measurement was taken at the most extreme point of the nucleus. For vowels without significant glides (e.g., (ah) and lax short *a*), the measurement was taken a point along the steady state. The sample rate was 10k Hz with the filter set at 4k Hz. Normalization of the vowel tokens was accomplished with a vowel normalization algorithm based on geometric mean (Neary 1977). t-tests and One-Way ANOVAs were performed on the means of the vowels under consideration to allow for statistical comparison of adults and children and of individual speakers. Vowel charts were created using Lotus 123, Release 5 (1991, 1994) and Freelance Graphics (1991,1993).

2.3 Variables

The variables examined in the current study included those vowels listed in Table 1. The variable (uw) was an overlapping variable in that it has been found to have undergone change in Philadelphia as well as being one of the vowels studied by Kuhl et al. (1997).

Table 1 Variables examined in children's and mothers' speech

Philadelphia Variables (Labov 1990)	
Nearly completed change	short *a*
Midrange changes	(ow), (uw)
New change	(aw), (eyC)
Triangle Variables (Kuhl et al. 1997)	(iy), (ah), (uw)

Specific research questions comprise the following:

1. How do the vowel productions of the children compare with those of their mothers in terms of their extremity or innovation?
2. What do the findings mean in terms of the children's role in their speech community and their effect on change in progress?
3. What are the possible mechanisms of acquisition and change that may help to account for the children's participation in the process of language variation and change?

3 Results

3.1 Philadelphia Vowels

Results of the study revealed that in all cases the speakers vowels overlapped substantially. Two cases of significant difference were found between adults and children. In addition, individual differences were noted among the children which support the importance of input to acquisition of dialect features.

3.1.1 Comparison of mothers and children

For all of the vowels measured, the children's productions overlapped those of their mothers. There were no significant differences in the F1 and F2 measurements for short *a*, (aw), and (uw). In the case of long (ow), the mean F2 measurement of the adult vowels was significantly greater than that for the children (t(93) = 4.32, $p = .004$) indicating that the adults productions of (ow) are more fronted than the children's productions.

The reversal of the lowering of checked, long (ey), referred to as (eyC) raising, is the one instance in which the children were producing more extreme forms than their care takers (F1, t(93) = 2.5, $p = .014$; F2, t(93) = 4.24, $p = 000$). This is a finding that begs further study since it is one of the newer changes, and it does not appear to be stigmatized. This lack of stigma is important since it diminishes the possibility of the most obvious initial interpretation, that the adults are style shifting during the interview. The style shifting explanation is initially appealing because it would explain not only the adult's more conservative productions but also the children's more extreme ones, as children have been found in other research not to style shift (Roberts 1997b). However, as (eyC) raising does not appear to be associated with reduced prestige, the alternate explanation, that the children are in fact producing more extreme tokens of this variable, is possible.

3.1.2 Comparison of Individual Children

Analysis of the children's individual data was accomplished by graphing of their tokens for each of the variables in question. In addition, One-way ANOVAs with Tukey HSD post hoc analyses, with the significance level set at ($p < .05$) were run to compare individual speaker means for the critical variables. That is, comparisons were made for F2 when the critical dialect feature was fronting (e.g., (uw) and (ow)) and for F1 and F2 when both raising and tensing or fronting were relevant (e.g., (eyC), (aw), and short *a*). The individual Ns are admittedly small, and the results of the statistical comparisons must be seen as preliminary. Nevertheless, the qualitative and quantitative analyses, taken together, both support the importance of early input to preschool dialect acquisition and resultant effects on change and provide direction for future research.

First of all, the children of native Philadelphia parents – Jenny, Danny, Evan, and Shelley – form a cohesive group. For only one of the five vowels do they differ significantly. For the fronting of (aw), Danny's productions are significantly more fronted than those of all of the other children. There is more individual variation when the two children with non-native parents are considered, however. The three simplest vowel changes, the fronting of (uw) and (ow), and the fronting and raising (aw), will be considered first. Roberts (1997a) found that when the rate of fronting was considered, these changes were among the easiest for these children to acquire. These comparisons of vowel extremity bear out these findings. There were no significant differences for any of the children in regards to (uw) fronting. Although Mike's tokens consistently among the less extreme, there was no significant difference between his mean and those of the others. In the fronting of (ow), there were also no significant differences, with the interesting exception that Gia's mean F2 measurement was significantly higher than that of the two lowest means, Mike's and Evan's. Mike's and Evan's productions did not differ significantly from those of the others, however. Finally, for (aw), in addition to the previously mentioned more fronted productions of Danny, the only significant difference for F2 was that of Mike, whose productions were less extreme than all of the other children's but Shelley's. For F1, there were several significant individual differences among the children. Most importantly, however, was the finding that Mike's tokens were significantly lower than those of Jenny, Danny, Evan, and Gia, whose productions were generally the highest. In summary, these simpler changes were produced similarly by all of the children, even those with non-native parents. Gia's productions were indistinguishable from the others for (uw) and (aw), and more extreme for (ow). Mike, whose parents are from Italy, generally showed less extreme productions, although this difference was not significant for (uw) and (ow).

The raising of (eyC) can be seen as a more complex change, since it is blocked by a word boundary, but not a syllable boundary. Roberts (1997a) found that it was more difficult for a child with less Philadelphia input to learn. Gia, however, exhibits no significant differences from the other children, although her tokens are generally clustered at the less extreme end of the range. Mike's productions, on the other hand, are significantly less extreme than the most extreme speakers in the group (Danny for F1; Danny and Jenny for F2), but not from those of the more midrange children. In other words, he is certainly acquiring this feature, but his productions are consistently among the least extreme in the group. Figure 1 illustrates this point with a vowel chart containing the individual (eyC) tokens of all of the children. Although there are more centralized tokens than those of Gia and Mike, these represent occasional tokens from several of the others, whose productions also included extreme tokens. Gia and Mike were responsible for none of the extreme tokens.

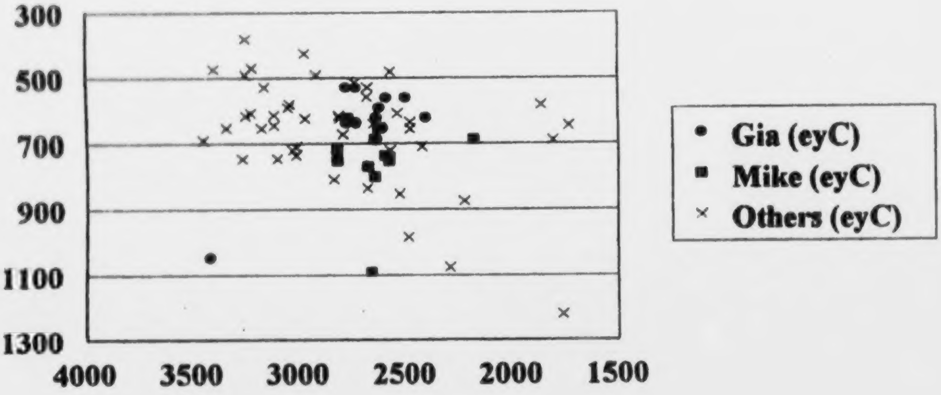


Figure 1: Raising of (eyC): Gia and Mike

Finally, short *a* is the most complex of the changes studied, with lexical, grammatical and phonological conditions affecting its occurrence. Not surprisingly, and in support of the findings of Roberts (1997a), the differences between the children with contrasting input situations shows up most for this variable. Significant differences for F1 between Mike and the two most extreme speakers (Danny and Jenny) show that his productions are among the least raised in the group. Similarly, significant differences between F1 and F2 for both Mike and Gia and the more extreme speakers (Jenny, Danny, and Evan for Mike; Jenny and Evan for Gia) reveal that their productions are also among the least raised.

Table 2 is provided as an illustration of the data analyzed and indicates the means, standard deviations, significance level, and numbers of tokens for the individual speakers for short *a*.

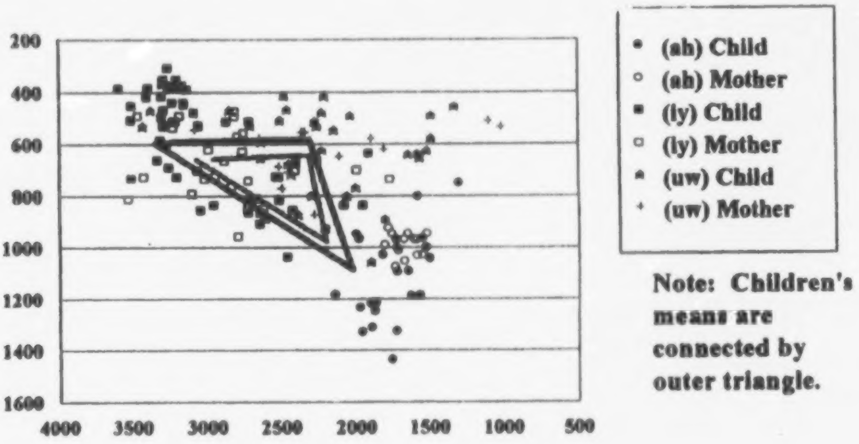
Variable	Speaker	N	Mean	S.D.	Significance
short *a* F1	Danny	19	639.5	89.9	NS
	Jenny	22	646.3	161.7	NS
	Evan	11	690.6	145.2	NS
	Shelley	12	715.8	152.7	NS
	Gia	14	817.4	331.2	NS
	Mike	10	897.7	281	$p < .05$
short *a* F2	Danny	19	2919.7	168.9	NS
	Jenny	22	3152.5	231.2	NS
	Evan	11	3169.5	509.1	NS
	Shelley	12	2854.4	169.5	NS
	Gia	14	2631	552	$p < .05$
	Mike	10	2535	211.9	$p < .05$

Table 2 Comparison of F1 and F2 means for short *a* for all children

In summary, these findings support those of Roberts (1997a). In addition to having more difficulty learning the more complex sound changes, the children with non-native parents also appeared to have generally less extreme productions. It is important to note, however, that their productions do overlap those of the other children. They are not out of "community range", merely on the less extreme end of it.

3.2 Triangle Point Vowels

The three vowels measured by Kuhl et al. (1997) and found to be more peripheral in the mothers' speech when talking to infants than when talking to adults were (iy), (uw), and (ah), described by the authors as the three points of the vowel triangle. These three vowels were also measured in the current study to ascertain whether the children's tokens were more extreme than those of their caretakers. If so, this could be seen as initial support for the idea that if the children are hearing more extreme tokens during CDS, these vowels may in fact become the targets of their future vowel productions. The results of this study showed that for (iy) and (ah) the children's productions were more extreme (i.e., toward the periphery of the vowel space) than those of their care takers. (See Figure 1 for and illustration of the following findings.) That is, for (iy), the children's mean F1 was significantly lower (t (73.36) = 2.4, *p* = .019) and their mean F2 was significantly higher (t (79) = 2.27, *p* = .026), resulting in productions that were more fronted than higher than those of their mothers.



Note: Children's means are connected by outer triangle.

Figure 2: Vowel triangle for children and mothers, (iy), (uw), (ah)

For (ah), the children's mean F1 was significantly higher than that of their mothers (t(39) =3.77, *p* = .001), but there was no significant difference between the F2 means resulting in a lowered production for the children. Long (uw), a vowel which also occurs as a Philadelphia variable, demonstrated no significant differences between care takers and children.[1]

4 Discussion

The results of the current study, in combination with those of previous work, underscore the role of young children as participants in and contributors to sound change in progress. Although the key role of adolescents in sound change has been well-documented, current research suggests that sound change does not begin in adolescence. The workings of dialect acquisition and change beginning in early childhood and extending into adolescence and beyond may provide fruitful avenues of research for some time to come. The purpose of this work has been to suggest possible domains in which this process can occur and potential learning mechanisms which could advance it to be explored in future research.

One mechanism considered by the current study is that of the input process of child-directed speech (CDS), familiar to researchers of language acquisition. The findings of Kuhl and colleagues are that during CDS, mothers' speech was characterized by a significantly expanded vowel triangle. They argued that this feature of CDS would aid the child in acquiring a phonological system by providing them more specific vowel information. The results of the current study revealed that the children's productions of two of the three triangle vowels, ((ah) and (iy), but not (uw)), were significantly more peripheral than those of their mothers. The possibility suggested by these findings is that the more peripheral input heard by the children in infancy may then become the targets for their own eventual productions. For the dialect variables in which an innovative variable is also the more peripheral variable (e.g., (eyC) and short *a* in Philadelphia), these early peripheral productions may become the typical productions of children as they mature, resulting in a more peripheral innovative form. A second, related possibility is that these early peripheral forms, in which children match the expanded

---

[1] To minimize the possibility that the normalization process had an effect on these results, means of short vowels (e.g., (I), (e)), in which no prediction of vowel triangle stretching would pertain, were made. There were no significant differences between adults and children in these results.

CDS productions, may become the domains in which adolescents, having experienced the expanded model, may further push the limits of the vowel space. The current findings, that children's productions of (iy) and (eyC) were more extreme than those of their mothers provide rudimentary support for these suggestions. Clearly, further research is needed, particularly into the question of what happens when the most innovative dialect form is *not* the most perepheral form, as in the case of the fronting of (ow) and (uw). Kuhl does not provide dialect background information on the speakers in her studies, so we do not know if this dialect situation would change the CDS productions of the mothers. The current study cannot answer this question either. The facts that (uw) was the only one of the triangle vowels not to result in more pheripheral productions by the children, and that the mothers produced more extreme tokens of (ow) than the children suggests that something different is happening in the back of the vowel space than in the front. Unfortunately, exactly what that something is will have to be the subject of future research. Two predictions to be explored at a later time, however, include the following: Children appear to be in a unique position to contribute to language change in the situations in which the more innovative form is also the more peripheral form. Secondly, like the situation in which the child care situation appears to lead to the favoring of female-led changes, it is predicted that the CDS situation could encourage a preference for changes involving vowel shifts toward the periphery of the vowel space.

The newness of this area of inquiry and its complexity leave us with many areas to explore. However, the results of previous and current research clearly converge on the point that children can and do contribute to change. Early indications from the current study are that this contribution may be aided by the same learning mechanisms that are relevant to language acquisition in general.

Like gender, ethnicity, and age, the construct of speech community is not fixed but socially negotiated and continually changing. Sociologists have argued that, rather than merely reflecting their identities as a function of biologically-influenced roles, adults and children construct, or 'do' race, ethnicity, gender, etc. I suggest that children, like adults, also 'do' community. Thorne (1993), in her research on the construction of gender by kindergarten children, argues "that within the complexities, within the 'play of gender', indeed lie possibilities of social change." By the same token, within the interactions in which children and adolescents participate, with their care takers, with their peers, with their community, lie the possibilities of language change.

References

Andruski, J. E. & Kuhl, P. K. 1996. The acoustic structure of vowels in mothers' speech to infants and adults. *Proceedings of the 1996 International Conference on Spoken Language Processing* , 3, 1545-1548.

Clarke-Stewart, K. 1973. Interactions between mothers and their young children: Characteristics and consequences. *Monographs of the Society for Research in Child Development, 38* (Serial No. 153).

Eckert, P. 1989. The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1: 245-268.

Eckert, P. 1988. Adolescent social structure and the spread of linguistic change. *Language in Society* 17: 183-207.

Freelance Graphics for Windows Release 2.01. 1991, 1993. Lotus Development Corporation.

Gleitman, L. and Wanner, E. 1982. Language acquisition: The state of the art. In E. Wanner and L. Gleitman (eds.) *Language acquisition: The state of the art.* Cambridge, MA: Harvard University Press.

Heath, S. B. 1983. *Ways with words.* Cambridge: Cambridge University Press.

Kerswill, P. 1996. Children, adolescents, and language change. *Language Variation and Change* 8: 177-202.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. 1997. Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277, 684-686.

Kuhl, P. K. & Meltzoff, A. N. 1996. Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustic Society of America*, 100, 2425-2438.

Labov, W. 1990. Sex and social class in linguistic change. *Language Variation and Change, 2*, 205-254.

Labov, W. 1984. Field methods on the Project on Linguistic Change and Variation. In Baugh, J. And J. Sherzer (eds.), *Language in use: Readings in sociolinguistics.* Englewood Cliffs, NJ: Prentice-Hall.

Lotus 1-2-3 Release 5 for Windows. 1991, 1994. Lotus Development Corporation.

Murray, A., Johnson, J. and Peters, J. 1990. Fine-tuning of utterance length to preverbal infants: Effects on later language development. *Journal of Child Language, 17*, 511-526.

Neary, T. 1977. Phonetic feature system for vowels. Unpublished dissertation, University of Connecticut.

Ochs, E. 1988. *Culture and language development.* Cambridge: Cambridge University Press.

BEST COPY AVAILABLE

Payne, A. 1990. Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In W. Labov (ed.) *Locating language in time and space*. New York: Academic.

Pinker, S. 1984. *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Roberts, J. 1997a. Hitting a moving target: Acquisition of sound change in progress by Philadelphia children. *Language Variation and Change, 9*, 249-266.

Roberts, J. 1997b. Acquisition of variable rules: A study of (-t,d) deletion in preschool children. *Journal of Child Language, 24*, 351-372.

Roberts, J. 1996. Acquisition of variable rules: A study of (-t,d) deletion in preschool children. IRCS Report 96-09. Philadelphia: Center for Research in Cognitive Science, University of Pennsylvania.

Roberts, J. & Labov, W. 1995. Learning to talk Philadelphian: Acquisition of short a by preschool children. *Language Variation and Change, 7*, 101-112.

Schieffelin, B. 1990. *The give and take of everyday life*. Cambridge: Cambridge University Press.

Van Ausdale, D. & Feagin, J. 1996. Using racial & ethnic concepts: The critical case of very young children. *American Sociological Review, 61*, 779-793.

Thorne, B. 1993. *Gender Play*. New Brunswick, NJ: Rutgers University Press.

West, C. & Zimmerman, D. 1995. Doing difference. *Gender & Society, 9*, 8-37.

West, C. & Zimmerman, D. 1987. Doing gender. *Gender & Society, 1*, 125-151.

Department of Communication Sciences
University of Vermont
Burlington, VT 05405
jroberts@polyglot.uvm.edu

Situated Ethnicities: Constructing and Reconstructing Identity in the Sociolinguistic Interview[*]

Natalie Schilling-Estes

1 Introduction

Over the past several decades, numerous quantitative variationist investigations have demonstrated how linguistic variables are used to indicate—and create—ethnic distinctiveness (e.g., Labov et al. 1968, Wolfram 1969, Labov 1972). However, traditional variationist studies are limited in that (1) they are often based on the assumption that individual ethnic identity is clear-cut and static, even though relations between ethnic groups may change over time, and (2) they do not show how individual speakers use language features in actual discourse to shape and project ethnic identity. Instead, they focus on the aggregate patterning of data which have been abstracted from the speech events in which they were originally situated.

Following researchers such as Bell (forthcoming, Bell and Johnson 1997), Rickford and McNair-Knox (1994), and the California Style Collective (Arnold et al. 1993), I present an analysis in which language features are investigated in the context in which they occur, in this case the sociolinguistic interview. I focus on one interview from a large-scale sociolinguistic study of Robeson County, North Carolina. This is a rural tri-ethnic community whose residents include Lumbee Native Americans, African Americans, and Anglo Americans (e.g., Wolfram 1996). The interview takes place between a Lumbee Native American and an African American from Wilmington, a small city on the North Carolina coast.

Using both qualitative and quantitative approaches, I examine the interview in terms of (1) usage levels of several ethnic and regional dialect features in different sections of the interview, (2) co-occurrence of features, and (3) strategic use of features during key moments. I also show the importance of discourse-level features in creating and marking alignments. The analysis reveals that ethnic identity is not static or monolithic but is shaped and reshaped on an on-going basis during conversational interaction. In addition, ethnicity is not the only component of individual and group identity, and speakers must balance considerations of ethnicity with such other matters as regional and family background and current social affiliations. And finally, even if we focus only on the ethnic component of identity, we find that speakers must perform balancing acts of various sorts—for example between locally defined and globally defined notions of ethnic group membership.

---

[*] This research was supported by NSF Grant No. SBR-96-16331. I would like to thank Walt Wolfram, Penelope Eckert, Janet Fuller, Scott Kiesling, and Keli Yerian for their help, support, and inspiration.

2 Community Background

Robeson County, located in Southeastern North Carolina, is home to approximately equal numbers of Whites, Blacks, and Native Americans (approximately 40,000 each). Apparently, the Lumbee have existed as a coherent people in this region since well before White and Black communities were firmly established in there. Their ancestral language is unknown, since the first records of Native Americans in the Robeson County area, in the early 1700s, describe an indigenous people who spoke English. Despite their early exposure to and apparent assimilation of White culture, the Lumbee have continually struggled to carve out a niche as a separate people who are neither White, Black, nor a mere amalgamation of better-known Indian tribes. Over the centuries, they have developed a distinctive dialect of English. This dialect is comprised of a few features not found in neighboring varieties (e.g., the use of *I'm* in present perfect contexts, as in *I'm forgot* for *I've forgotten*), as well as features which have been adapted from other varieties so that they pattern in unique ways (e.g., the use of *bees* rather than *be* in habitual contexts) (Dannenberg and Wolfram 1998).

3 The Interview

The interview that forms the focus of this analysis takes place between two young adult males, a Lumbee university student (the interviewee) and an African American (the interviewer) who attends the same university, has African American family connections in Robeson County, and self-identifies as part Cherokee Indian. The two met at the university a couple of years before the interview takes place, and they are good friends with many friends in common. The interview lasts approximately an hour and 15 minutes; I analyze all but the last few minutes, which are on a separate tape from the bulk of the interview.

I divided the interview into several sections based on topic. The interview begins with a discussion of race relations which can be broken into several subsections: race relations in Robeson County in general (7 minutes, 8 seconds), race relations in the county during the Civil War (3:15), and race-related issues in current politics (1:46). Following the discussion of race relations, the two interlocutors move on to a relatively brief discussion of two of the Lumbee's family members: his brother (1:04), and his uncle (3:17). They then turn to a lengthy discussion of mutual friends at the university. Twenty minutes and 45 seconds later, they abruptly resume their discussion of race relations. This time, the discussion encompasses the following subtopics: race relations in Robeson County (2:55), race relations during the Civil War (2:57), race relations in the South in general (2:16), and race relations on a national and global level (11:31).

I examined the patterning of a number of phonological and morphosyntactic features in each section and subsection of the interview, including (1) postvocalic *r*-lessness, as in *fahm* for *farm* or *cah* for *car*; (2) regularization patterns for past tense *be*, as in *They wasn't there* or *She weren't home*; (3) third-person singular *–s* absence, as in *He walk to the store*; (4) habitual *be*, as in *John always be working late*; and (5) copula deletion, as in *He a nice guy*. Each of the variables I examined has either been the focus of a previous quantitative study of interethnic language difference in Robeson County or is a widespread and well-studied feature of African American Vernacular English. I hope that my case study will shed light on the meaning of the community-wide—and wider—patterns that have been observed in larger-scale analyses.

4 Phonological Variation: *R*-lessness

So far, it has not been easy to either describe or explain the interethnic patterning of *r*-lessness in Robeson County. Both Dannenberg 1998 and Miller 1996 have conducted quantitative analyses that show an intermediate level of *r*-lessness for Lumbee Native Americans vis-à-vis neighboring White and Black speakers. Further, both interpret this pattern as evidence that *r*-lessness serves as a marker of ethnic affiliation in Robeson County: Since the Lumbee consider themselves to be a separate people, their usage levels for *r*-lessness do not conform either to the relatively high levels exhibited by African Americans or the decreasing levels shown by Robeson County Whites. Despite the agreement between the two studies, each shows different alignment patterns among the groups over time, thus making it difficult to tell exactly what the variable really means to speakers in each ethnic group. Figures 1 and 2 illustrate.

The differing patterns for *r*-lessness uncovered in each study are due in part to the fact that the age groups in the two studies are not exactly parallel but more importantly to the fact that there is a great deal of heterogeneity in terms of *r*-lessness even within age cohorts. Thus, for example, speakers in

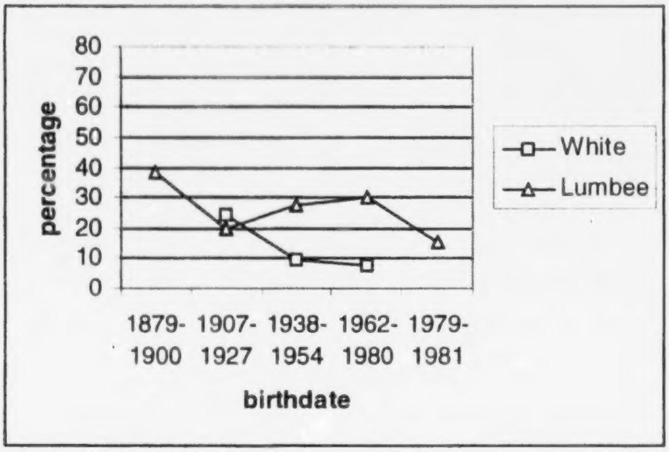


Figure 1: Community-wide patterns for *r*-lessness, adapted from Dannenberg 1998.

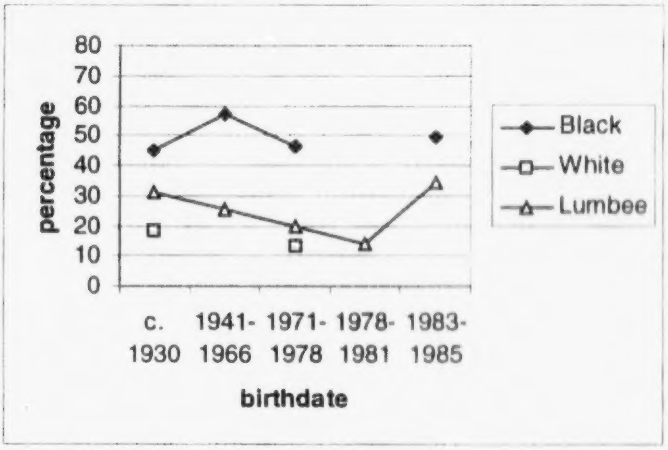


Figure 2: Community-wide patterns for *r*-lessness, adapted from Miller 1996

Dannenberg's oldest age group show levels of *r*-lessness from seven to 73 percent, while even those in the more homogeneous youngest group show levels from three to 29 percent.

The current analysis indicates that the heterogeneity also extends to individuals. This is illustrated in Figure 3, which shows the levels of *r*-lessness for each speaker in each section of the interview. For the purposes of this investigation, I will call the Lumbee speaker Dan and the African American Ronald.

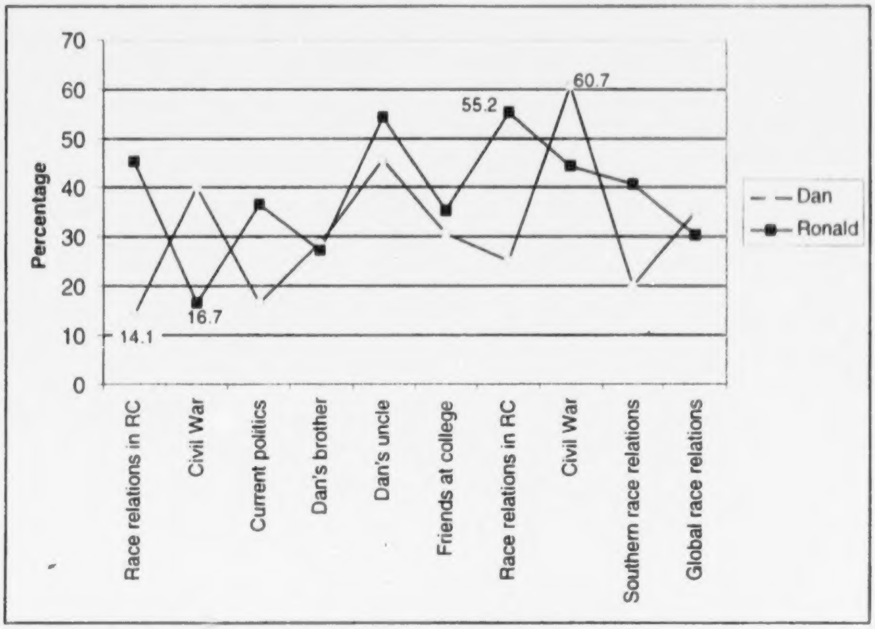


Figure 3: *R*-lessness in the individual interview

If we assume that *r*-lessness is indeed a marker of ethnic group membership in Robeson County, then we might expect to see Dan, the Lumbee, using consistently lower levels for *r*-lessness than Ronald, the African American, who will use the higher levels associated with AAVE across the country. However, as we can see, the two do not differ from each other by a consistent amount: They show considerably more divergence when talking about race relations than when talking about family and friends.

One explanation for this is that considerations of ethnic identity and ethnic group membership are more salient when the two are talking directly about this subject than when they are talking about Dan's family or about friends, at which point it becomes more important to indicate and forge personal bonds through linguistic convergence.[1] This explanation is borne out by the fact that the two are most convergent in the brief section about Dan's brother that occurs before the extended discussion of mutual friends. Even though the focus here is on Dan's family rather than common friends, considerations of personal friendship are extremely salient here. Ronald has abruptly shifted the topic from an argument about politics to a discussion of Dan's brother, who had been killed in a motorcycle accident a couple of years ago. Ronald had been under the impression that the accident took place quite a long time ago, and the discussion becomes awkward as he comes to the realization that he has inadvertently brought up a touchy subject. However, the two manage to quickly smooth over the awkwardness, even without overt apology from Ronald, partly through rapidly shifting away from the topic of Dan's brother to the topic of Dan's uncle, which they discuss in a humorous tone, and partly through their mutual convergence with respect to *r*-lessness.

---

[1] In Kiesling and Schilling-Estes (1998), we point out that linguistic convergence is not always indicative of psychological or social convergence with one's interlocutors, nor is linguistic divergence necessarily indicative of psychological divergence. For example, speakers who are hostile to one another may converge linguistically by hurling insults at one another, while a woman and man seeking increased intimacy may diverge linguistically by adopting prototypically "female" and "male" speech styles, respectively. In this, we follow recent versions of Communication Accommodation Theory (formerly Speech Accommodation Theory), as outlined, for example, in Giles, Coupland, and Coupland (1991). For the purposes

However, if we accept that levels of *r*-lessness are dependent on the relative importance of ethnic distinctiveness vs. personal connection in different portions of the interview, we are still left with the question of why the levels vary so much within the two sections on race relations and within the section on family and friends. I suggest that this is because *r*-lessness serves to index other meanings besides ethnic group membership. Not only is *r*-lessness associated with AAVE, but it is also a marker of Southern speech in general. Prior to WWII, it was a marker of prestigious speech; since that time, however, it has fallen out of social favor and now serves as a marker of vernacular Southern speech. Thus, Dan and Ronald increase their *r*-lessness when talking about Dan's uncle, a subject that pertains to the local vernacular culture. Conversely, they decrease their *r*-lessness when talking about their friends at the university, a non-local matter.

It is important to note here that just because the two speakers show comparatively low-levels of *r*-lessness in the section on friends does not mean that we should classify this section as more "formal" than the preceding one. The discussion of friends is highly animated, and the two interlocutors demonstrate a high degree of involvement with their subject matter and with each other, as indicated by convergence not only in terms of *r*-lessness but in terms of other variable features as well, as discussed below. Involvement is also indicated at the discourse level. For example, the two use discourse markers indicative of high involvement (e.g., "you know" and "I

---

of this discussion, I will remain, for the most part, within the traditional framework of SAT, which holds that speakers seeking psychological convergence will attempt linguistic convergence.

mean"; see Schiffrin 1987) to a greater extent in the section on friends than in the preceding sections. In addition, they directly address each other frequently in this section, even though there was no direct address in the initial section on race relations. Thus, it seems best to conclude that the section on friends is no less "informal" than the preceding section on Dan's uncle: Speakers simply exhibit a variety of speech styles when conversing informally and unguardedly, and it may be that the persistent sociolinguistic quest for each speaker's one true "vernacular" is somewhat misguided (cf. Wolfson 1976, Milroy 1987:57-60).

Because *r*-lessness has undergone such dramatic shifts in prestige value and usage levels in the past couple of generations, shifting to high levels of *r*-lessness carries other meanings besides membership in Southern vernacular communities: It also has the effect of evoking the past, including the speech of prestigious Southern Whites of days gone by. Thus, Dan dramatically increases his *r*-lessness each time he discusses the Civil War, especially in invented "quotes" from historic Southern whites and in his highly detailed and emotionally charged descriptions of Civil War heroes, such as his description of Robert E. Lee, transcribed in (1)[2].

(1) D: Man, he was dashing, you know. He had that black hair, he just rode around on his horse, [he was—he was bad.
R: [Uh huh.
D: By the end of the war, which only lasted four years, he looked o::ld, man.
R: He was old. He had gray—
D: [Beat do::wn.
R: [had gray—had gray—]
D: [Beat do::wn]
R: Had gray hair all on his face and stuff.
D: And he'd—he'd fought for so long. I mean, he really did.
R: [Uh huh.
D: [I mean, he really got involved.[3]

---

[2] Note that brackets indicate overlapping talk; parentheses indicate inserted explanatory material; hyphens indicate false starts; colons indicate extended vowel length; a series of periods indicates a pause (pauses were not timed for the purposes of this investigation).

[3] Dan's high levels of *r*-lessness in his discussion of the Civil War are perfect examples of cases in which linguistic divergence is not intended for—and does not lead to—psychological divergence. Dan is primarily evoking images through his *r*-lessness in these sections and hence heightening rather than diminishing the

We still have one final question with respect to *r*-lessness: Why do the two speakers show convergence at the end of the second section on race relations? This time, the explanation lies in the fact that style shifting with respect to variable dialect features depends on more than one's fellow conversational participants and the topic of conversation. It also depends on speakers' alignment toward one another and toward their topics. For example, throughout most of the second section on race relations, the two speakers not only discuss potentially divisive topics, but they exhibit overt disagreement—arguing over matters ranging from whether Native Americans are genetically predisposed toward alcoholism to whether such figures as Robert E. Lee, Malcolm X, and Martin Luther King, Jr. are worthy of admiration. Near the end of the section, though, they finally reach an agreement: Martin Luther King *is* to be admired, even though Ronald had initially doubted the effectiveness of his non-violent methods.

Even more important than the two interlocutors' alignment with respect to their subject matter, the two have also forged some crucial alignments with respect to one another. At various earlier points in the interview, Dan forges alliances between the Lumbee—and, by extension, himself—and Robeson County Whites and Whites in general. For example, at one point he states that Whites historically have treated the Lumbee differently than they treated Blacks and that the Lumbee "were just like White people, you know, they weren't subservient at all." On the other hand, Ronald is quite insistent, throughout the interview, on his dislike of and disassociation from Whites. For example, he maintains that "the White man will always find your weakness" and that Martin Luther King was taking a big risk indeed when he "gambled that White people had a conscience."

By the end of the second section on race relations, however, Dan has shifted alliances. Like Ronald, he now sides against Whites. For example, he sarcastically "quotes" generic Whites who extol the virtues of segregation by saying, "Why do you call this racism? This ain't nothing but, uh, segregation. Segregation, it works, it's good, it works for everybody." At the same time, Dan minimizes rather than heightens his personal distance from African Americans. He accomplishes this not only via overt statements (e.g., his praise for Martin Luther King) but through rejecting ties with the Lumbee, who historically have been antagonistic toward Blacks. In fact, the crucial argument that serves to convince Ronald that King's non-violent

---

interpersonal involvement between the two speakers. (See Tannen 1989, Chapter 6, on the role of imagery in creating conversational involvement.)

methods were effective after all hinges on the words of Dan's uncle, whose beliefs are diametrically opposed to Dan's. This passage is transcribed in (2).

(2) I'll tell you why they (White people) were scared of *him* (*King*). Because, uh ... He had brought so much change, and people so scared of change, they couldn't believe it. I had, I talked to my uncle, my—my dad's, my mom's uncle. Asked him, you know, what'd you think about Martin Luther King? And uh ... he said, "He's a son of a bitch!" He said, "He's a son of a bitch now, and—and they tried to build him up as something that he ain't, but he's a son of bitch now, and he was back then."

Interestingly, Dan indicates alignment with African Americans only after he and Ronald have shifted the focus of their discussion of race relations from a local to a national level—as if it is easier for Dan to identify with African Americans in the abstract that in a more personal way.

The depersonalization of the second section on race relations is evident not only in topic choice but in more subtle matters such as pronoun use: Although the two make frequent reference to the Lumbee, to Native Americans in general, and to African Americans, they never once refer to these groups with first or second person pronouns (e.g., Dan calls the Lumbee "Indians" or "they" rather than "we"). In addition, whereas Dan makes a couple oblique references to his personal Indian identity (for example, he says, "Indians are a very lone breed. I do myself, I don't stick with other Indians, because I'm, you know, I'm just a lone person."), Ronald makes no reference in this section to his own personal ethnic group membership at all—in sharp contrast to the initial section of the interview, where he makes a number of (unsuccessful) attempts to discuss his Robeson County roots.

Ronald further indicates disassociation from ethnic ties, particularly localized affiliations, through his usage levels for vernacular language features. Thus, in the final section on race relations, he shows 28.6% third person singular *–s* absence compared with 52.9% in the section on friends, and he shows only 12.5% copula deletion, compared with 55.6% in the first section on race relations and 45.8% in the discussion of friends. In addition, his use of *r*-lessness decreases steadily as he and Dan shift their focus from local to national matters.

Thus we see that even if we leave aside the other social meanings of *r*-lessness in Robeson County and focus solely on its role in indicating ethnic alignment, the picture is still highly complex, because ethnic alignments exist on a number of different levels, from the local to the regional to the global, and speakers may align themselves in various ways with respect to different types of ethnic group memberships during the course of a single conversation.

5 Morphosyntactic Features

Let us now turn to an examination of morphosyntactic features in this interview. Three of the features I examined may be considered to be markers of African American identity, since they are widespread in AAVE across the country. These are third person singular *-s* absence, copula deletion, and habitual *be*. In addition, past *be* regularization patterns are indicative of ethnic group membership in Robeson County. Regularization to *was* and *wasn't*, as in *They wasn't there*, is a widespread AAVE feature, while the Lumbee indicate a distinctive pattern of their own: regularization to *was* in affirmative contexts but to *weren't* in negative utterances, as is *They was home* vs. *He weren't home*. This pattern is discussed in detail in Wolfram and Sellers (forthcoming).

Overall, Dan uses hardly any of the morphosyntactic features associated with AAVE, and so it is meaningless to talk about differing usage levels in each different section of the interview. However, it is instructive to note that when Dan does use AAVE features, they are concentrated in the discussion of mutual friends, in which, as I have already discussed, considerations of interpersonal alignment are more important than ethnic distinctiveness. In addition, the positioning of these features within the discussion is important as well. For example, not only does one of Dan's two cases of third person singular *-s* absence occur in this section, but it echoes Ronald's use of the form, as illustrated in (3).

(3) D: Well, I mean.. I mean, he has found something that **he does** that I don't believe **nobody else does** better, and that's the secret to life.
R: Uh huh.
D: I mean, and that's the secret to a successful one.
R: Find something that **nobody else do**?
D: **Nobody else do** better.

Further, there are only three tokens of habitual *be* in the entire interview, and they all occur in quick succession, again in the section on friends. Interestingly, this time it is Dan rather than Ronald who introduces the form and Ronald who picks up on it. This is illustrated in (4), an excerpt from their discussion of a mutual friend who had joined a cult.

(4) D: It used to be a old, like, sixties kind of church? But they changed?
R: Uh huh.

D: Jack (term of address), they **be** telling them people some crazy stuff, Ron.
R: How you know that's the one Hunter's in?
D: 'Cause, 'cause, uh, Roger told me to watch it (a TV special on cults).
R: You joking. What's the name of it? You don't know what the name of it is? What—what they—what—what they **be** telling them?
D: They **be** telling them stuff like, uh, you got to get twenty members by the time you get in here. You get saved, then you got to get twenty members or you can't stay in this church.

Ronald also echoes Dan in terms of certain usages that are not associated with AAVE. For example, even though he shows his highest usage levels for third person singular *–s* absence in the section on friends, he still occasionally uses *–s* forms at certain strategic points. For example, in the excerpt in (5), he uses an *–s* form in completing Dan's thought; while in (6), he self-corrects to an *–s* form, indicating that he is working to overcome his propensity toward *–s* deletion in informal conversation in his efforts to converge with Dan.

(5) D: But you know, um, there's a old saying that goes.. Fondness—I mean, no, not fondness, no, separation? Being away from somebody?
R: **Makes** the heart [grow fonder.]
D: [grow fonder.]
R: Not for me.

(6) R: See, I don't know. That stuff (marijuana), it—it **mess** with you—it **messes** with your head so much.

Example (6) illustrates that when we're looking at so-called ethnic dialect features and what their usage levels might have to tell us about ethnic alignment, we have to keep in mind that linguistic convergence is not always easy—or possible. In other words, divergence or lack of convergence does not necessarily indicate that a speaker wishes to disassociate from an interlocutor but may simply indicate that the speaker lacks the linguistic ability to converge. For example, although Ronald is quite adept at manipulating his levels of *r*-lessness to converge with or diverge from Dan's, his ability to converge in terms of past *be* regularization is more limited. Overall, Dan shows limited regularization of past *be* in affirmative sentences (14%) but shows the classic Lumbee pattern in negative contexts—that is, regulariza-

BEST COPY AVAILABLE

tion to *weren't*, as *in He weren't home* or *It weren't me*. In fact, Dan never once uses *wasn't*, whether in regularized or non-regularized contexts (e.g., *They wasn't* vs. *he wasn't*). Ronald, on the other hand, mirrors the AAVE pattern in showing regularization to the *was-* form in both positive and negative contexts (*was* regularization = 30%; *wasn't* = 71.4%). There are only two cases in which Ronald uses *weren't* rather than *wasn't* for past *be*, and each of these occurs in a section in which considerations of friendship with Dan are uppermost. One case occurs in the middle of the discussion of friends, in a particularly animated section in which the two are talking about sex. This is illustrated in (7).

(7) R: Said Hunter was down there having sex with a girl on the couch in the study lounge. Anybody could walk in, [Dan!]
D: [laughter]
R: In the study lounge but that—but **you weren't**—**you wadn't** here when, uh, Jim and Jane—yeah **you was**.

Even though Ronald seems to be trying to converge with Dan, his self-“correction” from standard *you weren't* to nonstandard *you wasn't* indicates that he's having difficulty. And his switch back to his own system for past *be* regularization seems relatively permanent, since he uses *you was* right after he uses *you wasn't*.

The other case in which Ronald uses *weren't* is in the discussion of Dan's brother discussed above; a portion is transcribed in (8). This time, Ronald is a little more successful in converging with Dan, in that he uses *they weren't* without “correcting” to *they wasn't*; however, he returns to regularizing to *was* a few lines later.

(8) D And then—**they weren't** never the same after that.
R: **They weren't**?
D: Not after you lose [a child.
R: [They still—they've still changed? I mean, you can still, you still see they difference?
D: Yeah.
R: And that's been how many years now?
D: That's been, uh, seven, three, six years.
R: Dan, **that wadn't** too long ago. I thought **you was**—that was something happened when **you was** a little kid or something.
D: No::.

In addition, a few minutes later he makes another unsuccessful attempt to converge with Dan. This time he regularizes to the *were-* form but does so in an affirmative context; and further, he fails to regularize to *weren't* in an immediately following negative tag. This is shown in (9).

(9) Oh, **he were** beating on her or something, **wadn't he**?

We see, then, that when considering what speakers' usage patterns for morphosyntactic features have to tell us about their ethnic self- and group identification, it is important to investigate not only aggregate usage levels but also particular occurrences of important features in the discourse contexts in which they are situated. Further, we also have to keep in mind that morphosyntactic features may not be as easy to adopt as phonological features, and speakers may exhibit confusion between two differing systems despite their best efforts to converge with their fellow interlocutors.

6 Conclusion

In conclusion, I have attempted to demonstrate in this investigation the importance of recontextualizing the sociolinguistic interview data on which quantitative variationist studies are based. Through close examination of a range of ethnic language features in the discourse contexts in which they occur, we can arrive at a greater understanding of the full range of meanings associated with such features and of the range of meanings associated with the seemingly straightforward notion of ethnic group membership. We can also see that no matter how important ethnicity may be as a component of individual identity, particularly in interethnic encounters, it is only one of a myriad of such components, and speakers are far more adept at balancing innumerable considerations of individual identity and interpersonal relations than we often give them credit for. Speakers are not automatons whose use of variable features is nothing but a dull reflection of their demographic characteristics. Rather, they are active strategizers who use language to shape and re-shape themselves and the world around them. However, we can only get a full sense of speaker agency if we complement our large-scale quantitative studies with case studies that examine speech as it actually unfolds.

References

Arnold, Jennifer, et al. 1993. Variation and personal/group style. Paper presented at NWAV 22, Ottawa, Canada.

Bell, Allan. forthcoming. Back in style: Reworking audience design. In *Style and variation*, ed. John Rickford and Penelope Eckert. Cambridge/New York: Cambridge University Press.

Bell, Allan, and Johnson, Gary. 1997. Towards a sociolinguistics of style. In *University of Pennsylvania Working Papers in Linguistics 4.1: A Selection of Papers from NWAV 25*, 1-21. University of Pennsylvania.

Dannenberg, Clare J. 1998. Towards the understanding of ethnic distinction and *r*-lessness in multi-ethnic Southern communities: A study of Lumbee American Indian Vernacular English. Paper presented at SECOL 29, Lafayette, LA.

Dannenberg, Clare, and Walt Wolfram. 1998. Ethnic identity and grammatical restructuring: *Be(s)* in Lumbee English. *American Speech* 73, 139-159.

Giles, Howard, Nikolas Coupland, and Justine Coupland. 1991. Accommodation theory: Communication, context, and consequence. In *Contexts of accommodation: Developments in applied sociolinguistics*, ed. Howard Giles, Justine Coupland, and Nikolas Coupland, 1-68. Cambridge: Cambridge University Press.

Kiesling, Scott Fabius, and Natalie Schilling-Estes. 1998. Language Style as Identity Construction: A Footing and Framing Approach. Poster presented at NWAV 27, Athens, GA.

Labov, William. 1972. *Language in the inner city: Studies in the Black English vernacular*. Philadelphia: University of Pennsylvania Press.

Labov, William, Paul Cohen, Clarence Robins, and John Lewis. 1968. *A study of the non-standard English of Negro and Puerto Rican speakers in New York City*. United States Office of Education final Report, Research Project 3288.

Miller, Jason. 1996. *Mixed sociolinguistic alignment and ethnic identity: R-lessness in a Native American Community*. M.A. Thesis, North Carolina State University.

Milroy, Lesley. 1987. *Observing and analysing natural language*. Cambridge: Blackwell.

Rickford, John, and Faye McNair-Knox. 1994. Addressee- and topic-influenced style shift. In *Sociolinguistic perspectives on register*, ed. Douglas Biber and Edward Finegan, 235-76. New York: Oxford University Press.

Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.

Tannen, Deborah. 1989. *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.

Wolfram, Walt. 1969. *A linguistic description of Detroit Negro speech*. Washington, DC: Center for Applied Linguistics.

Wolfram, Walt. 1996. Delineation and description in dialectology: The case of perfective *I'm* in Lumbee English. *American Speech* 71, 5-26.

Wolfram, Walt, and Jason Sellers. forthcoming. Ethnolinguistic marking of past *be* in Lumbee Vernacular English. *Journal of English Linguistics*.

Wolfson, Nessa. 1976. Speech events and natural speech: Some implications for sociolinguistic methodology. *Language in Society* 5, 189-209.

Department of English
Old Dominion University
Batten Arts and Letters 220
Norfolk, VA 23529
nsestes@odu.edu

The *University of Pennsylvania Working Papers in Linguistics (PWPL)* is an occasional series produced by the Penn Linguistics Club, the graduate student organization of the Linguistics Department of the University of Pennsylvania.

Publication in this volume does not preclude submission of papers elsewhere; copyright is retained by the authors of the individual papers.

Volumes of the Working Papers are available for $15, prepaid. Please see our web site for additional information.

The PWPL Series Editors

Jim Alexander
Alexis Dimitriadis
Na-Rae Han
Elsi Kaiser
Michelle Minnick Fox
Christine Moisset
Alexander Williams

How to reach the PWPL

U. Penn Working Papers in Linguistics
Department of Linguistics
619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104–6305

http://www.ling.upenn.edu/papers/pwpl.html
working-papers@ling.upenn.edu

Contents of Volume 6 (1999)

Volume 6.1
Proceedings of the 23rd annual Penn Linguistics Colloquium

Eric Baković

Assimilation to the unmarked.

Marlyse Baptista

On the non-universality of functional projections and the effects on parametrized variation: Evidence from Creoles.

Natalie Batman-Ratyosyan and Karin Stromswold

What Turkish acquisition tells us about underlying word order and scrambling.

Emily Bender

Constituting context: Null objects in English recipes revisited.

Cedric Boeckx

Decomposing French questions.

Farida Cassimjee and Charles W. Kisseberth

A conspiracy argument for Optimality Theory: Emakhuwa dialectology.

Barbara Citko

Light-headed relatives.

Kristin M. Eide and Tor A. Åfarli

Semantically charged syntax and the construction of meaning.

Stefan Engelberg

'Punctuality' and verb semantics.

Miho Fujiwara.

Acquisition of default inflectional suffixes: Japanese adjectives.

Elena Gavruseva and Rosalind Thornton

Possessor extraction in child English: A Minimalist account.

Kook-Hee Gill

The long-distance anaphora conspiracy: The case of Korean.

JEANETTE K. GUNDEL
Topic, focus, and the grammar-pragmatics interface.

J.-R. HAYASHISHITA
Two ways of deriving distributive readings.

BEN HERMANS
The reduplicative nature of the Bulgarian definite article.

RONALD KIM
The origin of the Pre-Ossetic oblique case suffix and its implications.

ALEXEI KOCHETOV
A hierarchy of phonetic constraints on palatality in Russian.

WILLIAM MCCLURE
Japanese floating classifiers.

MARTHA MCGINNIS
A-scrambling exists!

MICHELLE MINNICK FOX AND KAZUAKI MAEDA
Perception and production of American English tense and lax vowels by Japanese speakers.

MASAO OCHI AND TIEN-HSIN HSIN
The syntax of adjunct wh-NPs.

KIERAN SNYDER
Reflexives and constraints on the borrowing of discourse function: Creoles and Tahitian French.

MIKAEL VINKA
Two types of verb particle constructions.

Volume 6.2
Selected papers from NWAV(E) 27

JULIE AUGER AND JEFFREY STEELE
Vowel epenthesis in Vimeu Picard: A preliminary investigation.

Pamela Grant-Russell and Céline Beaudet
Lexical borrowings from French in written Quebec English.

Hikyoung Lee
Variable article use in Korean Learners of English.

Mimi Lipson
The loss of auxiliary selection in English.

Rajend Mesthrie
Syntactic change in progress: Semi-auxiliary *busy* in South African English.

Miriam Meyerhoff
The emergence of creole subject-verb agreement.

Naomi Nagy and Hélène Blondeau
Double subject marking in L2 Montreal French.

Peter Patrick
Testing the creole continuum.

Julie Roberts
Going younger to do difference: The role of children in language change.

Natalie Schilling-Estes
Situated ethnicities: Constructing and reconstructing identity in the sociolinguistic interview.

Volume 6.3
Current work in linguistics

Jennifer Arnold
The effects of thematic roles on pronoun use and frequency of reference.

Steven Bird
Multidimensional exploration of online linguistic field data.

Eugene Buckley
Extrametricality in Latin and Manam.

Robin Clark and Gerhard Jäger

A categorial syntax for verbs of perception.

Chunghye Han

On negative alternative questions.

Al Kim

The convergence of lexicalist perspectives in psycholinguistics and computational linguistics.

Ronald Kim

The distribution of the Old Irish infixed pronouns: Evidence for the syntactic evolution of Insular Celtic?

Eon-suk Ko

The laryngeal effect in Korean: Phonology or phonetics?

Seth Kulick

Defective complements in Tree Adjoining Grammar.

Jeff Lidz

Echo reduplication in Kannada: Implications for a theory of word formation.

Martha McGinnis

Event heads and the distribution of Psych-Verbs.

Rashmi Prasad and Michael Strube

Pronominal reference in Hindi: An incremental account.

Ellen Prince and Susan Pintzuk

Bilingual code-switching and the open/closed class distinction.

Amanda Seidl

Yoruba vowel elision and compounding.

Clarissa Surek-Clark

Dialect acquisition and prestige.

Order Form

U. Penn Working Papers in Linguistics

Subscription: $35 US for three issues, beginning with volume 7.1

Single issues: $15 US each (except vol. 4.1)

Please see our web site for complete tables of contents.

Order	#	Price	
SUBSCRIPTION (VOLUMES 7.1+)	___	@ $35	= ________
VOLUME 1, 1994 Working Papers	___	@ $15	= ________
VOLUME 2.1, 1995 Working Papers................	___	@ $15	= ________
VOLUME 2.2, Proceedings of PLC 19	___	@ $15	= ________
VOLUME 3.1, Papers from NWAVE 24	___	@ $15	= ________
VOLUME 3.2, Proceedings of PLC 20	___	@ $15	= ________
VOLUME 4.1, Papers from NWAVE 25	___	**@ $18**	= ________
VOLUME 4.2, Proceedings of PLC 21..............	___	@ $15	= ________
VOLUME 4.3, 1997 Working Papers	___	@ $15	= ________
VOLUME 5.1, Proceedings of PLC 22	___	@ $15	= ________
VOLUME 5.2 1998 Working Papers................	___	@ $15	= ________
VOLUME 6.1, THIS VOLUME	___	@ $15	= ________
VOLUME 6.2 Papers from NWAVE 27...........	___	@ $15	= ________
VOLUME 6.3, 1999 Working Papers	___	@ $15	= ________

Check to *Penn Linguistics Club*, in the amount of: ________

To place an order, please:

Detach this page by cutting along left margin (or copy both sides).
Give your address on the back of this form.
Enclose payment (check or money order).

Mail completed form and payment to:

PWPL
619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104–6305

Questions? *working-papers@babel.ling.upenn.edu*
Or consult our web page: *http://www.ling.upenn.edu/papers/pwpl.html*

(Please do not order volumes by email; we need to receive payment in advance.)

✂

Name: ______________________________

Address: ______________________________

____

____

____

Email: ______________________________

University of Pennsylvania Working Papers in Linguistics

Volume 6.2 (1999)
Selected Papers from NWAV(E) 27

JULIE AUGER AND JEFFREY STEELE. Vowel epenthesis in Vimeu Picard: A preliminary investigation 1

PAMELA GRANT-RUSSELL AND CÉLINE BEAUDET. Lexical borrowings from French in written Quebec English 17

HIKYOUNG LEE. Variable article use in Korean Learners of English 35

MIMI LIPSON. The loss of auxiliary selection in English 49

RAJEND MESTHRIE. Syntactic change in progress: Semi-auxiliary *busy* in South African English 63

MIRIAM MEYERHOFF. The emergence of creole subject-verb agreement 77

NAOMI NAGY AND HÉLÈNE BLONDEAU. Double subject marking in L2 Montreal French 93

PETER PATRICK. Testing the creole continuum 109

JULIE ROBERTS. Going younger to do difference: The role of children in language change 123

NATALIE SCHILLING-ESTES. Situated ethnicities: Constructing and reconstructing identity in the sociolinguistic interview 139

ISSN 1524-9549